

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

**(19) World Intellectual Property Organization
International Bureau**



(43) International Publication Date
20 March 2003 (20.03.2003)

PCT

(10) International Publication Number
WO 03/023008 A2

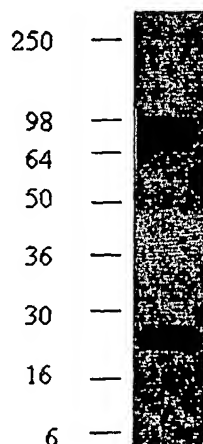
(51) International Patent Classification⁷:		C12N	60/379,532	10 May 2002 (10.05.2002)	US
(21) International Application Number:		PCT/US02/28596	60/381,672	17 May 2002 (17.05.2002)	US
(22) International Filing Date:			60/381,664	17 May 2002 (17.05.2002)	US
			60/383,651	28 May 2002 (28.05.2002)	US
		9 September 2002 (09.09.2002)	60/384,012	29 May 2002 (29.05.2002)	US
			60/390,155	19 June 2002 (19.06.2002)	US
			Not furnished	6 September 2002 (06.09.2002)	US
(25) Filing Language:		English	(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:		
(26) Publication Language:		English			
(30) Priority Data:			US	60/390,155 (CIP)	
			Filed on	19 June 2002 (19.06.2002)	
60/318,120	7 September 2001 (07.09.2001)	US	(71) Applicant: CURAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th floor, New Haven, CT 06511 (US).		
60/318,219	7 September 2001 (07.09.2001)	US			
60/318,130	7 September 2001 (07.09.2001)	US	(72) Inventors; and		
60/318,430	10 September 2001 (10.09.2001)	US			
60/318,765	12 September 2001 (12.09.2001)	US	(75) Inventors/Applicants (for US only): ZHONG, Mei [CA/US]; 45 Harrison Avenue, Apartment 1B, Branford, CT 06405 (US). LI, Li [CN/US]; 56 Jerimoth Drive, Branford, CT 06405 (US). GORMAN, Linda [US/US]; 329 Monticello Drive, Branford, CT 06405 (US). SPYTEK, Kimberly, A. [US/US]; 28 Court Street, number 1, New Haven, CT 06511 (US). KEKUDA, Ramesh [IN/US]; 71 Aiken Street, Unit R3, Norwalk, CT 06851 (US). TAUPIER, Raymond, J., Jr. [US/US]; 34 Pardee Place Extension, East Haven, CT 06512 (US). ANDERSON, David, W. [US/US]; 85 Montoya Drive, Branford, CT 06405 (US). VERNET, Corine, A., M. [FR/US]; 1739 Foxon Road, Apartment L6, Branford, CT 06471		
60/322,781	17 September 2001 (17.09.2001)	US			
60/322,816	17 September 2001 (17.09.2001)	US			
60/323,519	19 September 2001 (19.09.2001)	US			
60/323,631	20 September 2001 (20.09.2001)	US			
60/323,636	20 September 2001 (20.09.2001)	US			
60/325,091	25 September 2001 (25.09.2001)	US			
60/324,969	25 September 2001 (25.09.2001)	US			
60/324,990	26 September 2001 (26.09.2001)	US			
60/357,303	15 February 2002 (15.02.2002)	US			
60/360,973	28 February 2002 (28.02.2002)	US			
60/366,131	20 March 2002 (20.03.2002)	US			
60/367,753	25 March 2002 (25.03.2002)	US			
60/369,479	2 April 2002 (02.04.2002)	US			

[Continued on next page]

(54) Title: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS OF USE

NOV30b (CG51117-05) protein secreted by 293 cells.

Mw (kDa)



(57) Abstract: Disclosed herein are nucleic acid sequences that encode novel polypeptides. Also disclosed are polypeptides encoded by these nucleic acid sequences, and antibodies that immunospecifically bind to the polypeptide, as well as derivatives, variants, mutants, or fragments of the novel polypeptide, polynucleotide, or antibody specific to the polypeptide. Vectors, host cells, antibodies and recombinant methods for producing the polypeptides and polynucleotides, as well as methods for using same are also included. The invention further discloses therapeutic, diagnostic and research methods for diagnosis, treatment, and prevention of disorders involving any one of these novel human nucleic acids and proteins.

WO 03/023008 A2



(US). **CATTERTON, Elina** [FI/US]; 584, Boston Post Road, Madison, CT 06443 (US). **MILLER, Charles, E.** [US/US]; 98 Saddle Hill Drive, Guilford, CT 06437 (US). **SHENOY, Suresh, G.** [IN/US]; 15 Millwood Drive, Branford, CT 06405 (US). **PATTURAJAN, Meera** [IN/US]; 45 Harrison Avenue, Apartment 1C, Branford, CT 06405 (US). **PENA, Carol, E., A.** [US/US]; 604 Orange Street, Number 2, New Haven, CT 06511 (US). **TCHERNEV, Velizar, T.** [BG/US]; 45 Jefferson Road, #3-12, Branford, CT 063405 (US). **PADIGARU, Muralidhara** [IN/US]; 71 Hampton Park, Branford, CT 06405 (US). **GUSEV, Vladimir, Y.** [UA/US]; 1209 Durham Road, Madison, CT 06443 (US). **MALYANKAR, Uriel, M.** [IN/US]; 229 Branford Road, number 330, Branford, CT 06405 (US). **BURGESS, Catherine, E.** [US/US]; 90 Carriage Hill Drive, Wethersfield, CT 06109 (US). **GERLACH, Valerie, L.** [US/US]; 18 Rock Pasture Road, Branford, CT 06405 (US). **CASMAN, Stacie, J.** [US/US]; 17 Peck Street, North Haven, CT 06473 (US). **RIEGER, Daniel, K.** [DE/US]; 10A McKinnel Court, Branford, CT 06405 (US). **GROSSE, William, M.** [US/US]; 15 Rice Terrace Road, Apartment C, Branford, CT 06405 (US). **SMITHSON, Glenda** [US/US]; 125 Michael Drive, Guilford, CT 06435 (US). **PEYMAN, John, A.** [US/US]; 336 West Rock Avenue, New Haven, CT 06515 (US). **STARLING, Gary** [NZ/US]; 45 Robin Court, Middletown, CT 06457 (US). **ROTHENBERG, Mark, E.** [US/US]; 2 Allen Road, Clinton, CT 06413 (US). **LAROCHELLE, William, J.** [US/US]; 15 Devonshire Lane, Madison, CT 06443 (US). **SHIMKETS, Richard, A.** [US/US]; 5 Indian Meadows Drive, Guilford, CT 06437 (US). **CRABTREE, Julie** [US/US]; 426 NorthWest 19th Avenue, Gainesville, FL 32609 (US). **RASTELLI, Luca** [IT/US]; 52 Pepperbush Lane, Guilford, CT 06437 (US). **VOSS, Edward, Z.** [US/US]; 123 Knollwood Drive, Wallingford, CT 06492 (US). **BOLDOG, Ferenc, L.** [HU/US]; 1687 Hartford Turnpike, North Haven, CT 06473 (US). **EDINGER, Shlomit, R.** [US/US]; 766 Edgewood Avenue, New Haven, CT 06515 (US). **MILLET,**

Isabelle [FR/US]; 74 Carrington Avenue, Milford, CT 06460 (US). **MACDOUGALL, John, R.** [US/US]; 117 Russel Street, Hamden, CT 06517 (US). **ELLERMAN, Karen** [US/US]; 87 Montoya Drive, Branford, CT 06405 (US). **CHAPOVAL, Andrei** [US/US]; 16 Commercial Street, Branford, CT 06405 (US).

(74) **Agent: ELRIFI, Ivor, R.;** Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo, P., C., One Financial Center, Boston, MA 02111 (US).

(81) **Designated States (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS OF USE

FIELD OF THE INVENTION

The present invention relates to novel polypeptides, and the nucleic acids encoding them, having properties related to stimulation of biochemical or physiological responses in a cell, a tissue, an organ or an organism. More particularly, the novel polypeptides are gene products of novel genes, or are specified biologically active fragments or derivatives thereof. Methods of use encompass diagnostic and prognostic assay procedures as well as methods of treating diverse pathological conditions.

BACKGROUND OF THE INVENTION

Eukaryotic cells are characterized by biochemical and physiological processes which under normal conditions are exquisitely balanced to achieve the preservation and propagation of the cells. When such cells are components of multicellular organisms such as vertebrates, or more particularly organisms such as mammals, the regulation of the biochemical and physiological processes involves intricate signaling pathways. Frequently, such signaling pathways involve extracellular signaling proteins, cellular receptors that bind the signaling proteins, and signal transducing components located within the cells.

Signaling proteins may be classified as endocrine effectors, paracrine effectors or autocrine effectors. Endocrine effectors are signaling molecules secreted by a given organ into the circulatory system, which are then transported to a distant target organ or tissue. The target cells include the receptors for the endocrine effector, and when the endocrine effector binds, a signaling cascade is induced. Paracrine effectors involve secreting cells and receptor cells in close proximity to each other, for example two different classes of cells in the same tissue or organ. One class of cells secretes the paracrine effector, which then reaches the second class of cells, for example by diffusion through the extracellular fluid. The second class of cells contains the receptors for the paracrine effector; binding of the effector results in induction of the signaling cascade that elicits the corresponding biochemical or physiological effect. Autocrine effectors are highly analogous to paracrine effectors, except that the same cell type that secretes the autocrine effector also contains the receptor. Thus the autocrine effector binds to receptors on the same cell, or on identical neighboring cells. The binding process then elicits the characteristic biochemical or physiological effect.

Signaling processes may elicit a variety of effects on cells and tissues including by way of nonlimiting example induction of cell or tissue proliferation, suppression of growth or proliferation, induction of differentiation or maturation of a cell or tissue, and suppression of differentiation or maturation of a cell or tissue.

Many pathological conditions involve dysregulation of expression of important effector proteins. In certain classes of pathologies the dysregulation is manifested as diminished or suppressed level of synthesis and secretion of protein effectors. In other classes of pathologies the dysregulation is manifested as increased or up-regulated

level of synthesis and secretion of protein effectors. In a clinical setting a subject may be suspected of suffering from a condition brought on by altered or mis-regulated levels of a protein effector of interest. Therefore there is a need to assay for the level of the protein effector of interest in a biological sample from such a subject, and to compare the level with
5 that characteristic of a nonpathological condition. There also is a need to provide the protein effector as a product of manufacture. Administration of the effector to a subject in need thereof is useful in treatment of the pathological condition. Accordingly, there is a need for a method of treatment of a pathological condition brought on by a diminished or suppressed levels of the protein effector of interest. In addition, there is a need for a method
10 of treatment of a pathological condition brought on by a increased or up-regulated levels of the protein effector of interest.

Antibodies are multichain proteins that bind specifically to a given antigen, and bind poorly, or not at all, to substances deemed not to be cognate antigens. Antibodies are comprised of two short chains termed light chains and two long chains termed heavy chains.
15 These chains are constituted of immunoglobulin domains, of which generally there are two classes: one variable domain per chain, one constant domain in light chains, and three or more constant domains in heavy chains. The antigen-specific portion of the immunoglobulin molecules resides in the variable domains; the variable domains of one light chain and one heavy chain associate with each other to generate the antigen-binding
20 moiety. Antibodies that bind immunospecifically to a cognate or target antigen bind with high affinities. Accordingly, they are useful in assaying specifically for the presence of the antigen in a sample. In addition, they have the potential of inactivating the activity of the antigen.

Therefore there is a need to assay for the level of a protein effector of interest in a
25 biological sample from such a subject, and to compare this level with that characteristic of a nonpathological condition. In particular, there is a need for such an assay based on the use of an antibody that binds immunospecifically to the antigen. There further is a need to inhibit the activity of the protein effector in cases where a pathological condition arises from elevated or excessive levels of the effector based on the use of an antibody that binds
30 immunospecifically to the effector. Thus, there is a need for the antibody as a product of manufacture. There further is a need for a method of treatment of a pathological condition

brought on by an elevated or excessive level of the protein effector of interest based on administering the antibody to the subject.

SUMMARY OF THE INVENTION

The invention is based in part upon the discovery of isolated polypeptides including
5 amino acid sequences selected from mature forms of the amino acid sequences selected
from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127.
The novel nucleic acids and polypeptides are referred to herein as NOVX, or NOV1,
NOV2, NOV3, *etc.*, nucleic acids and polypeptides. These nucleic acids and polypeptides,
as well as derivatives, homologs, analogs and fragments thereof, will hereinafter be
10 collectively designated as "NOVX" nucleic acid or polypeptide sequences.

The invention also is based in part upon variants of a mature form of the amino acid
sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer
between 1 and 127, wherein any amino acid in the mature form is changed to a different
amino acid, provided that no more than 15% of the amino acid residues in the sequence of
15 the mature form are so changed. In another embodiment, the invention includes the amino
acid sequences selected from the group consisting of SEQ ID NO:2n, wherein n is an
integer between 1 and 127. In another embodiment, the invention also comprises variants of
the amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is
an integer between 1 and 127, wherein any amino acid specified in the chosen sequence is
20 changed to a different amino acid, provided that no more than 15% of the amino acid
residues in the sequence are so changed. The invention also involves fragments of any of
the mature forms of the amino acid sequences selected from the group consisting of SEQ ID
NO:2n, wherein n is an integer between 1 and 127, or any other amino acid sequence
selected from this group. The invention also comprises fragments from these groups in
25 which up to 15% of the residues are changed.

In another embodiment, the invention encompasses polypeptides that are naturally
occurring allelic variants of the sequence selected from the group consisting of SEQ ID
NO:2n, wherein n is an integer between 1 and 127. These allelic variants include amino acid
sequences that are the translations of nucleic acid sequences differing by a single nucleotide
30 from nucleic acid sequences selected from the group consisting of SEQ ID NOS: 2n-1,

wherein n is an integer between 1 and 127. The variant polypeptide where any amino acid changed in the chosen sequence is changed to provide a conservative substitution.

In another embodiment, the invention comprises a pharmaceutical composition involving a polypeptide with an amino acid sequence selected from the group consisting of
5 SEQ ID NO:2 n , wherein n is an integer between 1 and 127, and a pharmaceutically acceptable carrier. In another embodiment, the invention involves a kit, including, in one or more containers, this pharmaceutical composition.

In another embodiment, the invention includes the use of a therapeutic in the manufacture of a medicament for treating a syndrome associated with a human disease, the
10 disease being selected from a pathology associated with a polypeptide with an amino acid sequence selected from the group consisting of SEQ ID NO:2 n , wherein n is an integer between 1 and 127, wherein said therapeutic is the polypeptide selected from this group.

In another embodiment, the invention comprises a method for determining the presence or amount of a polypeptide with an amino acid sequence selected from the group
15 consisting of SEQ ID NO:2 n , wherein n is an integer between 1 and 127, in a sample, the method involving providing the sample; introducing the sample to an antibody that binds immunospecifically to the polypeptide; and determining the presence or amount of antibody bound to the polypeptide, thereby determining the presence or amount of polypeptide in the sample.

20 In another embodiment, the invention includes a method for determining the presence of or predisposition to a disease associated with altered levels of a polypeptide with an amino acid sequence selected from the group consisting of SEQ ID NO:2 n , wherein n is an integer between 1 and 127, in a first mammalian subject, the method involving measuring the level of expression of the polypeptide in a sample from the first mammalian
25 subject; and comparing the amount of the polypeptide in this sample to the amount of the polypeptide present in a control sample from a second mammalian subject known not to have, or not to be predisposed to, the disease, wherein an alteration in the expression level of the polypeptide in the first subject as compared to the control sample indicates the presence of or predisposition to the disease.

30 In another embodiment, the invention involves a method of identifying an agent that binds to a polypeptide with an amino acid sequence selected from the group consisting of

SEQ ID NO:2n, wherein n is an integer between 1 and 127, the method including introducing the polypeptide to the agent; and determining whether the agent binds to the polypeptide. The agent could be a cellular receptor or a downstream effector.

5 In another embodiment, the invention involves a method for identifying a potential therapeutic agent for use in treatment of a pathology, wherein the pathology is related to aberrant expression or aberrant physiological interactions of a polypeptide with an amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127, the method including providing a cell expressing the polypeptide of the invention and having a property or function ascribable to the polypeptide; contacting the
10 cell with a composition comprising a candidate substance; and determining whether the substance alters the property or function ascribable to the polypeptide; whereby, if an alteration observed in the presence of the substance is not observed when the cell is contacted with a composition devoid of the substance, the substance is identified as a potential therapeutic agent.

15 In another embodiment, the invention involves a method for screening for a modulator of activity or of latency or predisposition to a pathology associated with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127, the method including administering a test compound to a test animal at increased risk for a pathology associated with the
20 polypeptide of the invention, wherein the test animal recombinantly expresses the polypeptide of the invention; measuring the activity of the polypeptide in the test animal after administering the test compound; and comparing the activity of the protein in the test animal with the activity of the polypeptide in a control animal not administered the polypeptide, wherein a change in the activity of the polypeptide in the test animal relative to
25 the control animal indicates the test compound is a modulator of latency of, or predisposition to, a pathology associated with the polypeptide of the invention. The recombinant test animal could express a test protein transgene or express the transgene under the control of a promoter at an increased level relative to a wild-type test animal. The promoter may or may not be the native gene promoter of the transgene.

30 In another embodiment, the invention involves a method for modulating the activity of a polypeptide with an amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127, the method including introducing a cell

sample expressing the polypeptide with a compound that binds to the polypeptide in an amount sufficient to modulate the activity of the polypeptide.

In another embodiment, the invention involves a method of treating or preventing a pathology associated with a polypeptide with an amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127, the method including administering the polypeptide to a subject in which such treatment or prevention is desired in an amount sufficient to treat or prevent the pathology in the subject. The subject could be human.

In another embodiment, the invention involves a method of treating a pathological state in a mammal, the method including administering to the mammal a polypeptide in an amount that is sufficient to alleviate the pathological state, wherein the polypeptide is a polypeptide having an amino acid sequence at least 95% identical to a polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127, or a biologically active fragment thereof.

In another embodiment, the invention involves an isolated nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide having an amino acid sequence selected from the group consisting of a mature form of the amino acid sequence given SEQ ID NO:2n, wherein n is an integer between 1 and 127; a variant of a mature form of the amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127, wherein any amino acid in the mature form of the chosen sequence is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence of the mature form are so changed; the amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127; a variant of the amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127, in which any amino acid specified in the chosen sequence is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence are so changed; a nucleic acid fragment encoding at least a portion of a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 127, or any variant of the polypeptide wherein any amino acid of the chosen sequence is changed to a different amino acid, provided that no more than 10% of the amino acid residues in the sequence are so changed; and the complement of any of the nucleic acid molecules.

In another embodiment, the invention comprises an isolated nucleic acid molecule having a nucleic acid sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a mature form of the amino acid sequence given SEQ ID NO:2n, wherein n is an integer between 1 and 127, wherein the nucleic acid molecule
5 comprises the nucleotide sequence of a naturally occurring allelic nucleic acid variant.

In another embodiment, the invention involves an isolated nucleic acid molecule including a nucleic acid sequence encoding a polypeptide having an amino acid sequence selected from the group consisting of a mature form of the amino acid sequence given SEQ ID NO:2n, wherein n is an integer between 1 and 127, that encodes a variant polypeptide,
10 wherein the variant polypeptide has the polypeptide sequence of a naturally occurring polypeptide variant.

In another embodiment, the invention comprises an isolated nucleic acid molecule having a nucleic acid sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a mature form of the amino acid sequence given SEQ
15 ID NO:2n, wherein n is an integer between 1 and 127, wherein the nucleic acid molecule differs by a single nucleotide from a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 2n-1, wherein n is an integer between 1 and 127.

In another embodiment, the invention includes an isolated nucleic acid molecule having a nucleic acid sequence encoding a polypeptide including an amino acid sequence
20 selected from the group consisting of a mature form of the amino acid sequence given SEQ ID NO:2n, wherein n is an integer between 1 and 127, wherein the nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of the nucleotide sequence selected from the group consisting of SEQ ID NO:2n-1, wherein n is an integer between 1 and 127; a nucleotide sequence wherein one or more nucleotides in the
25 nucleotide sequence selected from the group consisting of SEQ ID NO:2n-1, wherein n is an integer between 1 and 127, is changed from that selected from the group consisting of the chosen sequence to a different nucleotide provided that no more than 15% of the nucleotides are so changed; a nucleic acid fragment of the sequence selected from the group consisting of SEQ ID NO:2n-1, wherein n is an integer between 1 and 127; and a nucleic
30 acid fragment wherein one or more nucleotides in the nucleotide sequence selected from the group consisting of SEQ ID NO:2n-1, wherein n is an integer between 1 and 127, is

changed from that selected from the group consisting of the chosen sequence to a different nucleotide provided that no more than 15% of the nucleotides are so changed.

In another embodiment, the invention includes an isolated nucleic acid molecule having a nucleic acid sequence encoding a polypeptide including an amino acid sequence selected from the group consisting of a mature form of the amino acid sequence given SEQ ID NO:2n, wherein n is an integer between 1 and 127, wherein the nucleic acid molecule hybridizes under stringent conditions to the nucleotide sequence selected from the group consisting of SEQ ID NO:2n-1, wherein n is an integer between 1 and 127, or a complement of the nucleotide sequence.

In another embodiment, the invention includes an isolated nucleic acid molecule having a nucleic acid sequence encoding a polypeptide including an amino acid sequence selected from the group consisting of a mature form of the amino acid sequence given SEQ ID NO:2n, wherein n is an integer between 1 and 127, wherein the nucleic acid molecule has a nucleotide sequence in which any nucleotide specified in the coding sequence of the chosen nucleotide sequence is changed from that selected from the group consisting of the chosen sequence to a different nucleotide provided that no more than 15% of the nucleotides in the chosen coding sequence are so changed, an isolated second polynucleotide that is a complement of the first polynucleotide, or a fragment of any of them.

In another embodiment, the invention includes a vector involving the nucleic acid molecule having a nucleic acid sequence encoding a polypeptide including an amino acid sequence selected from the group consisting of a mature form of the amino acid sequence given SEQ ID NO:2n, wherein n is an integer between 1 and 127. This vector can have a promoter operably linked to the nucleic acid molecule. This vector can be located within a cell.

In another embodiment, the invention involves a method for determining the presence or amount of a nucleic acid molecule having a nucleic acid sequence encoding a polypeptide including an amino acid sequence selected from the group consisting of a mature form of the amino acid sequence given SEQ ID NO:2n, wherein n is an integer between 1 and 127, in a sample, the method including providing the sample; introducing the sample to a probe that binds to the nucleic acid molecule; and determining the presence or amount of the probe bound to the nucleic acid molecule, thereby determining the presence

or amount of the nucleic acid molecule in the sample. The presence or amount of the nucleic acid molecule is used as a marker for cell or tissue type. The cell type can be cancerous.

In another embodiment, the invention involves a method for determining the presence of or predisposition for a disease associated with altered levels of a nucleic acid molecule having a nucleic acid sequence encoding a polypeptide including an amino acid sequence selected from the group consisting of a mature form of the amino acid sequence given SEQ ID NO:2n, wherein n is an integer between 1 and 127, in a first mammalian subject, the method including measuring the amount of the nucleic acid in a sample from the first mammalian subject; and comparing the amount of the nucleic acid in the sample of step (a) to the amount of the nucleic acid present in a control sample from a second mammalian subject known not to have or not be predisposed to, the disease; wherein an alteration in the level of the nucleic acid in the first subject as compared to the control sample indicates the presence of or predisposition to the disease.

The invention further provides an antibody that binds immunospecifically to a NOVX polypeptide. The NOVX antibody may be monoclonal, humanized, or a fully human antibody. Preferably, the antibody has a dissociation constant for the binding of the NOVX polypeptide to the antibody less than 1×10^{-9} M. More preferably, the NOVX antibody neutralizes the activity of the NOVX polypeptide.

In a further aspect, the invention provides for the use of a therapeutic in the manufacture of a medicament for treating a syndrome associated with a human disease, associated with a NOVX polypeptide. Preferably the therapeutic is a NOVX antibody.

In yet a further aspect, the invention provides a method of treating or preventing a NOVX-associated disorder, a method of treating a pathological state in a mammal, and a method of treating or preventing a pathology associated with a polypeptide by administering a NOVX antibody to a subject in an amount sufficient to treat or prevent the disorder.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of

conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

5

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a Western blot showing expression of NOV30b (CG51117-05) immunoreactive polypeptide in human embryonic kidney 293 cells.

Figure 2 is a schematic diagram of the x-ray crystal structure of porcine colipase and tetra ethylene glycol monooctyl ether inhibitor.

10 Figure 3 is a schematic diagram showing the interfacial binding domain of colipase.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides novel nucleotides and polypeptides encoded thereby. Included in the invention are the novel nucleic acid sequences, their encoded polypeptides, antibodies, and other related compounds. The sequences are collectively referred to herein
15 as "NOVX nucleic acids" or "NOVX polynucleotides" and the corresponding encoded polypeptides are referred to as "NOVX polypeptides" or "NOVX proteins." Unless indicated otherwise, "NOVX" is meant to refer to any of the novel sequences disclosed herein. Table A provides a summary of the NOVX nucleic acids and their encoded polypeptides.

20

TABLE A. Sequences and Corresponding SEQ ID Numbers

NOVX Assignme nt	Internal Identification	SEQ ID NO (nucleic acid)	SEQ ID NO (amino acid)	Homology
NOV1a	CG108440-01	1	2	Fibronectin precursor protein-like protein
NOV1b	CG108440-02	3	4	Fibronectin precursor protein-like protein
NOV2a	CG122589-01	5	6	Asialoglycoprotein receptor 2-like protein
NOV2b	CG122589-02	7	8	Asialoglycoprotein receptor 2-like protein
NOV2c	CG122589-03	9	10	Asialoglycoprotein receptor 2-like protein

NOV3a	CG133274-01	11	12	Induced myeloid leukemia cell differentiation protein MCL-1-like protein
NOV3b	CG133274-02	13	14	Induced myeloid leukemia cell differentiation protein MCL-1-like protein
NOV3c	278876765	15	16	Induced myeloid leukemia cell differentiation protein MCL-1-like protein
NOV3d	278881214	17	18	Induced myeloid leukemia cell differentiation protein MCL-1-like protein
NOV4a	CG134430-01	19	20	RIKEN cDNA 2310034104-like protein
NOV5a	CG137677-01	21	22	RIKEN 5730409G15-like protein
NOV6a	CG137697-01	23	24	RIKEN 5730409G15-like protein
NOV7a	CG137717-01	25	26	FLJ37712 fis protein-like protein
NOV8a	CG137793-01	27	28	High affinity immunoglobulin epsilon receptor alpha subunit precursor protein-like protein
NOV8b	CG137793-02	29	30	High affinity immunoglobulin epsilon receptor alpha subunit precursor protein-like protein
NOV9a	CG137873-01	31	32	Fibrinogen alpha chain precursor protein-like protein
NOV9b	CG137873-03	33	34	Fibrinogen alpha chain precursor protein-like protein
NOV9c	CG137873-02	35	36	Fibrinogen alpha chain precursor protein-like protein
NOV10a	CG137882-01	37	38	FLJ21269-like protein
NOV10b	CG137882-02	39	40	FLJ21269-like protein
NOV11a	CG137910-01	41	42	FLJ21432-like protein
NOV12a	CG138013-01	43	44	Sialic acid-binding immunoglobulin-like lectin-9-like protein
NOV13a	CG138074-01	45	46	RIKEN 2310012P03-like protein
NOV14a	CG138573-01	47	48	Folate receptor 3-like protein
NOV15a	CG138606-01	49	50	Brush border 61.9 KDa protein precursor-like protein

NOV16a	CGI38751-01	51	52	cAMP inducible 2 protein-like protein
NOV16b	CGI38751-02	53	54	cAMP inducible 2 protein-like protein
NOV17a	CGI39062-01	55	56	Jagged 1 precursor protein-like protein
NOV17b	CGI39062-02	57	58	Jagged 1 precursor protein-like protein
NOV18a	CGI39363-01	59	60	Transmembrane protein HTMP10-like protein
NOV18b	CGI39363-02	61	62	Transmembrane protein HTMP10-like protein
NOV19a	CGI40188-01	63	64	DC2-like protein
NOV20a	CGI40305-01	65	66	Complement-clq tumor necrosis factor-related protein-like protein
NOV20b	CGI40305-02	67	68	Complement-clq tumor necrosis factor-related protein-like protein
NOV21a	CGI40639-01	69	70	Flotillin-2 (Reggie-1) (REG-1)-like protein
NOV21b	CGI40639-02	71	72	Flotillin-2 (Reggie-1) (REG-1)-like protein
NOV22a	CGI40843-01	73	74	Integrin beta-5 precursor protein-like protein
NOV23a	CGI41540-01	75	76	IL1 receptor-type 2-like protein
NOV23b	CGI41540-02	77	78	IL1 receptor-type 2-like protein
NOV24a	CGI41580-01	79	80	KIAA 1467 protein-like protein
NOV25a	CGI41643-01	81	82	RIKEN 2010001CC9 protein-like protein
NOV26a	CGI42003-01	83	84	Plasma protease C1 inhibitor precursor protein-like protein
NOV26b	306076006	85	86	Plasma protease C1 inhibitor precursor protein-like protein
NOV26c	278889088	87	88	Plasma protease C1 inhibitor precursor protein-like protein
NOV26d	CGI42003-02	89	90	Plasma protease C1 inhibitor precursor protein-like protein
NOV27a	CGI42023-01	91	92	6230421J19Rik protein-like protein
NOV28a	CGI42092-01	93	94	C4b-binding protein alpha chain precursor protein-like protein
NOV28b	CGI42092-02	95	96	C4b-binding protein alpha chain precursor protein-like protein
NOV28c	CGI42092-03	97	98	C4b-binding protein alpha chain precursor protein-like protein

NOV29a	CG171681-01	99	100	Sushi repeat-containing protein
NOV29b	CG171681-03	101	102	Sushi repeat-containing protein
NOV29c	CG171681-02	103	104	Sushi repeat-containing protein
NOV30a	CG51117-01	105	106	Nephronectin-like protein
NOV30b	CG51117-05	107	108	Nephronectin-like protein
NOV30c	CG51117-06	109	110	Nephronectin-like protein
NOV30d	CG51117-07	111	112	Nephronectin-like protein
NOV30e	CG51117-03	113	114	Nephronectin-like protein
NOV30f	CG51117-02	115	116	Nephronectin-like protein
NOV30g	CG51117-04	117	118	Nephronectin-like protein
NOV30h	CG51117-08	119	120	Nephronectin-like protein
NOV30i	CG51117-09	121	122	Nephronectin-like protein
NOV31a	CG51264-01	123	124	ST7-like protein
NOV31b	CG51264-03	125	126	ST7-like protein
NOV31c	CG51264-04	127	128	ST7-like protein
NOV31d	CG51264-06	129	130	ST7-like protein
NOV31e	CG51264-07	131	132	ST7-like protein
NOV31f	CG51264-02	133	134	ST7-like protein
NOV31g	CG51264-05	135	136	ST7-like protein
NOV31h	CG51264-08	137	138	ST7-like protein
NOV31i	CG51264-09	139	140	ST7-like protein
NOV31j	CG51264-10	141	142	ST7-like protein
NOV31k	CG51264-11	143	144	ST7-like protein
NOV31l	CG51264-12	145	146	ST7-like protein
NOV31m	CG51264-13	147	148	ST7-like protein
NOV31n	CG51264-14	149	150	ST7-like protein
NOV31o	CG51264-15	151	152	ST7-like protein
NOV31p	CG51264-16	153	154	ST7-like protein
NOV32a	CG52423-01	155	156	PV-1-like protein
NOV33a	CG52919-01	157	158	Sez-6-like protein
NOV33b	CG52919-02	159	160	Sez-6-like protein
NOV33c	CG52919-03	161	162	Sez-6-like protein
NOV33d	CG52919-04	163	164	Sez-6-like protein
NOV33e	CG52919-05	165	166	Sez-6-like protein
NOV33f	CG52919-06	167	168	Sez-6-like protein
NOV33g	CG52919-01	169	170	Sez-6-like protein
NOV33h	CG52919-07	171	172	Sez-6-like protein
NOV33i	CG52919-08	173	174	Sez-6-like protein
NOV33j	CG52919-09	175	176	Sez-6-like protein
NOV34a	CG55698-01	177	178	Colipase precursor protein-like protein
NOV34b	CG55698-02	179	180	Colipase precursor protein-like protein
NOV35a	CG55832-01	181	182	Tenascin-C precursor protein-like protein

NOV35b	CG55832-03	183	184	Tenascin-C precursor protein-like protein
NOV35c	CG55832-02	185	186	Tenascin-C precursor protein-like protein
NOV36a	CG56054-01	187	188	Integrin alpha 7-like protein
NOV36b	CG56054-03	189	190	Integrin alpha 7-like protein
NOV36c	CG56054-04	191	192	Integrin alpha 7-like protein
NOV36d	CG56054-05	193	194	Integrin alpha 7-like protein
NOV36e	CG56054-06	195	196	Integrin alpha 7-like protein
NOV36f	CG56054-07	197	198	Integrin alpha 7-like protein
NOV36g	CG56054-08	199	200	Integrin alpha 7-like protein
NOV36h	CG56054-09	201	202	Integrin alpha 7-like protein
NOV36i	CG56054-10	203	204	Integrin alpha 7-like protein
NOV36j	CG56054-11	205	206	Integrin alpha 7-like protein
NOV36k	CG56054-12	207	208	Integrin alpha 7-like protein
NOV36l	CG56054-13	209	210	Integrin alpha 7-like protein
NOV36m	CG56054-14	211	212	Integrin alpha 7-like protein
NOV36n	CG56054-15	213	214	Integrin alpha 7-like protein
NOV36o	CG56054-16	215	216	Integrin alpha 7-like protein
NOV36p	CG56054-17	217	218	Integrin alpha 7-like protein
NOV36q	CG56054-18	219	220	Integrin alpha 7-like protein
NOV36r	CG56054-19	221	222	Integrin alpha 7-like protein
NOV36s	CG56054-02	223	224	Integrin alpha 7-like protein
NOV37a	CG88634-01	225	226	KIAA1219-like protein
NOV38a	CG97012-01	227	228	Seizure 6 precursor protein-like protein
NOV38b	CG97012-02	229	230	Seizure 6 precursor protein-like protein
NOV38c	CG97012-03	231	232	Seizure 6 precursor protein-like protein
NOV38d	CG97012-01	233	234	Seizure 6 precursor protein-like protein
NOV38e	210120300	235	236	Seizure 6 precursor protein-like protein
NOV38f	210120376	237	238	Seizure 6 precursor protein-like protein
NOV38g	210120463	239	240	Seizure 6 precursor protein-like protein
NOV38h	210120269	241	242	Seizure 6 precursor protein-like protein
NOV38i	CG97012-04	243	244	Seizure 6 precursor protein-like protein
NOV38j	CG97012-05	245	246	Seizure 6 precursor protein-like protein
NOV39a	CG99754-01	247	248	RIKEN protein-like protein
NOV39b	CG99754-02	249	250	RIKEN protein-like protein
NOV40a	CG99777-01	251	252	CD30 ligand-like protein
NOV40b	CG99777-02	253	254	CD30 ligand-like protein

Table A indicates the homology of NOVX polypeptides to known protein families. Thus, the nucleic acids and polypeptides, antibodies and related compounds according to the invention corresponding to a NOVX as identified in column 1 of Table A will be useful
 5 in therapeutic and diagnostic applications implicated in, for example, pathologies and disorders associated with the known protein families identified in column 5 of Table A.

Pathologies, diseases, disorders, conditions, and the like that are associated with NOVX sequences include, but are not limited to: *e.g.*, cardiomyopathy, atherosclerosis, hypertension, congenital heart defects, aortic stenosis, atrial septal defect (ASD),
 10 atrioventricular (A-V) canal defect, ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberous sclerosis, scleroderma, obesity, metabolic disturbances associated with obesity, adrenoleukodystrophy, congenital adrenal hyperplasia, prostate cancer, diabetes, metabolic disorders, neoplasm, hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies, graft versus
 15 host disease, AIDS, bronchial asthma, Crohn's disease; multiple sclerosis, treatment of Albright Hereditary Osteodystrophy, infectious disease, anorexia, cancer-associated cachexia, , neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disease, immune disorders, hematopoietic disorders, and the various dyslipidemias, the metabolic syndrome X, wasting disorders associated with chronic diseases, cancer, *e.g.*, uterine cancer,
 20 lymphoma, adenocarcinoma, as well as conditions such as transplantation, neuroprotection, fertility, or regeneration (*in vitro* and *in vivo*).

NOVX nucleic acids and their encoded polypeptides are useful in a variety of applications and contexts. The various NOVX nucleic acids and polypeptides according to the invention are useful as novel members of the protein families according to the presence
 25 of domains and sequence relatedness to previously described proteins. Additionally, NOVX nucleic acids and polypeptides can also be used to identify proteins that are members of the family to which the NOVX polypeptides belong.

Consistent with other known members of the family of proteins, identified in column 5 of Table A, the NOVX polypeptides of the present invention show homology to,
 30 and contain domains that are characteristic of, other members of such protein families. Details of the sequence relatedness and domain analysis for each NOVX are presented in Example A.

The NOVX nucleic acids and polypeptides can also be used to screen for molecules, which inhibit or enhance NOVX activity or function. Specifically, the nucleic acids and polypeptides according to the invention may be used as targets for the identification of small molecules that modulate or inhibit diseases associated with the protein families listed in Table A.

The NOVX nucleic acids and polypeptides are also useful for detecting specific cell types. Details of the expression analysis for each NOVX are presented in Example C. Accordingly, the NOVX nucleic acids, polypeptides, antibodies and related compounds according to the invention will have diagnostic and therapeutic applications in the detection of a variety of diseases with differential expression in normal vs. diseased tissues, *e.g.* detection of a variety of cancers.

Additional utilities for NOVX nucleic acids and polypeptides according to the invention are disclosed herein.

NOVX clones

NOVX nucleic acids and their encoded polypeptides are useful in a variety of applications and contexts. The various NOVX nucleic acids and polypeptides according to the invention are useful as novel members of the protein families according to the presence of domains and sequence relatedness to previously described proteins. Additionally, NOVX nucleic acids and polypeptides can also be used to identify proteins that are members of the family to which the NOVX polypeptides belong.

The NOVX genes and their corresponding encoded proteins are useful for preventing, treating or ameliorating medical conditions, *e.g.*, by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the NOVX genes, based on the tissues in which they are most highly expressed. Uses include developing products for the diagnosis or treatment of a variety of diseases and disorders.

The NOVX nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications and as a research tool. These include serving as a specific or selective nucleic acid or protein diagnostic and/or prognostic marker, wherein the presence or amount of the nucleic acid or the protein are to be assessed, as well as

potential therapeutic applications such as the following: (i) a protein therapeutic, (ii) a small molecule drug target, (iii) an antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) a nucleic acid useful in gene therapy (gene delivery/gene ablation), and (v) a composition promoting tissue regeneration *in vitro* and *in vivo* (vi) a biological defense weapon.

In one specific embodiment, the invention includes an isolated polypeptide comprising an amino acid sequence selected from the group consisting of: (a) a mature form of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 127; (b) a variant of a mature form of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 127, wherein any amino acid in the mature form is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence of the mature form are so changed; (c) an amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 127; (d) a variant of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 127, wherein any amino acid specified in the chosen sequence is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence are so changed; and (e) a fragment of any of (a) through (d).

In another specific embodiment, the invention includes an isolated nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of: (a) a mature form of the amino acid sequence given SEQ ID NO: 2n, wherein n is an integer between 1 and 127; (b) a variant of a mature form of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 127, wherein any amino acid in the mature form of the chosen sequence is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence of the mature form are so changed; (c) the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 127; (d) a variant of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 127, in which any amino acid specified in the chosen sequence is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence are so changed; (e) a nucleic acid fragment encoding at least a portion of a polypeptide comprising the

amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 127, or any variant of said polypeptide wherein any amino acid of the chosen sequence is changed to a different amino acid, provided that no more than 10% of the amino acid residues in the sequence are so changed; and (f) the complement of any of
5 said nucleic acid molecules.

In yet another specific embodiment, the invention includes an isolated nucleic acid molecule, wherein said nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of: (a) the nucleotide sequence selected from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 127; (b) a
10 nucleotide sequence wherein one or more nucleotides in the nucleotide sequence selected from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 127 is changed from that selected from the group consisting of the chosen sequence to a different nucleotide provided that no more than 15% of the nucleotides are so changed; (c) a nucleic acid fragment of the sequence selected from the group consisting of SEQ ID NO:
15 2n-1, wherein n is an integer between 1 and 127; and (d) a nucleic acid fragment wherein one or more nucleotides in the nucleotide sequence selected from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 127, is changed from that selected from the group consisting of the chosen sequence to a different nucleotide provided that no more than 15% of the nucleotides are so changed.

20 **NOVX Nucleic Acids and Polypeptides**

One aspect of the invention pertains to isolated nucleic acid molecules that encode NOVX polypeptides or biologically active portions thereof. Also included in the invention are nucleic acid fragments sufficient for use as hybridization probes to identify NOVX-encoding nucleic acids (*e.g.*, NOVX mRNAs) and fragments for use as PCR
25 primers for the amplification and/or mutation of NOVX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule may be single-stranded or double-stranded, but preferably is comprised
30 double-stranded DNA.

A NOVX nucleic acid can encode a mature NOVX polypeptide. As used herein, a "mature" form of a polypeptide or protein disclosed in the present invention is the product of a naturally occurring polypeptide or precursor form or proprotein. The naturally occurring polypeptide, precursor or proprotein includes, by way of nonlimiting example, the full-length gene product encoded by the corresponding gene. Alternatively, it may be defined as the polypeptide, precursor or proprotein encoded by an ORF described herein. The product "mature" form arises, by way of nonlimiting example, as a result of one or more naturally occurring processing steps that may take place within the cell (*e.g.*, host cell) in which the gene product arises. Examples of such processing steps leading to a "mature" form of a polypeptide or protein include the cleavage of the N-terminal methionine residue encoded by the initiation codon of an ORF, or the proteolytic cleavage of a signal peptide or leader sequence. Thus a mature form arising from a precursor polypeptide or protein that has residues 1 to N, where residue 1 is the N-terminal methionine, would have residues 2 through N remaining after removal of the N-terminal methionine. Alternatively, a mature form arising from a precursor polypeptide or protein having residues 1 to N, in which an N-terminal signal sequence from residue 1 to residue M is cleaved, would have the residues from residue M+1 to residue N remaining. Further as used herein, a "mature" form of a polypeptide or protein may arise from a step of post-translational modification other than a proteolytic cleavage event. Such additional processes include, by way of non-limiting example, glycosylation, myristylation or phosphorylation. In general, a mature polypeptide or protein may result from the operation of only one of these processes, or a combination of any of them.

The term "probe", as utilized herein, refers to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), about 100 nt, or as many as approximately, *e.g.*, 6,000 nt, depending upon the specific use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are generally obtained from a natural or recombinant source, are highly specific, and much slower to hybridize than shorter-length oligomer probes. Probes may be single-stranded or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

The term "isolated" nucleic acid molecule, as used herein, is a nucleic acid that is separated from other nucleic acid molecules which are present in the natural source of the

nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5'- and 3'-termini of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated NOVX nucleic acid molecules can contain less than
5 about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell/tissue from which the nucleic acid is derived (*e.g.*, brain, heart, liver, spleen, *etc.*). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium, or of chemical precursors or other chemicals.

10 A nucleic acid molecule of the invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, or a complement of this nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, as a
15 hybridization probe, NOVX molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, *et al.*, (eds.), MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

20 A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template with appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to NOVX nucleotide sequences can be
25 prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides
30 comprise a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment of the invention, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at

least 6 contiguous nucleotides of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, or a complement thereof. Oligonucleotides may be chemically synthesized and may also be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises
5 a nucleic acid molecule that is a complement of the nucleotide sequence shown in SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, or a portion of this nucleotide sequence (*e.g.*, a fragment that can be used as a probe or primer or a fragment encoding a biologically-active portion of a NOVX polypeptide). A nucleic acid molecule that is complementary to the nucleotide sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer
10 between 1 and 127, is one that is sufficiently complementary to the nucleotide sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, that it can hydrogen bond with few or no mismatches to the nucleotide sequence shown in SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, thereby forming a stable duplex.

As used herein, the term “complementary” refers to Watson-Crick or Hoogsteen
15 base pairing between nucleotides units of a nucleic acid molecule, and the term “binding” means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, van der Waals, hydrophobic interactions, and the like. A physical interaction can be either direct or indirect. Indirect interactions may be through or due to the effects of
20 another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

A “fragment” provided herein is defined as a sequence of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific
25 hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, and is at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice.

A full-length NOVX clone is identified as containing an ATG translation start codon
30 and an in-frame stop codon. Any disclosed NOVX nucleotide sequence lacking an ATG start codon therefore encodes a truncated C-terminal fragment of the respective NOVX polypeptide, and requires that the corresponding full-length cDNA extend in the 5' direction

of the disclosed sequence. Any disclosed NOVX nucleotide sequence lacking an in-frame stop codon similarly encodes a truncated N-terminal fragment of the respective NOVX polypeptide, and requires that the corresponding full-length cDNA extend in the 3' direction of the disclosed sequence.

5 A “derivative” is a nucleic acid sequence or amino acid sequence formed from the native compounds either directly, by modification or partial substitution. An “analog” is a nucleic acid sequence or amino acid sequence that has a structure similar to, but not identical to, the native compound, *e.g.* they differs from it in respect to certain components or side chains. Analogs may be synthetic or derived from a different evolutionary origin
10 and may have a similar or opposite metabolic activity compared to wild type. A “homolog” is a nucleic acid sequence or amino acid sequence of a particular gene that is derived from different species.

Derivatives and analogs may be full length or other than full length. Derivatives or analogs of the nucleic acids or proteins of the invention include, but are not limited to,
15 molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the invention, in various embodiments, by at least about 70%, 80%, or 95% identity (with a preferred identity of 80-95%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of
20 hybridizing to the complement of a sequence encoding the proteins under stringent, moderately stringent, or low stringent conditions. *See e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below.

A “homologous nucleic acid sequence” or “homologous amino acid sequence,” or
25 variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences include those sequences coding for isoforms of NOVX polypeptides. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the invention,
30 homologous nucleotide sequences include nucleotide sequences encoding for a NOVX polypeptide of species other than humans, including, but not limited to: vertebrates, and thus can include, *e.g.*, frog, mouse, rat, rabbit, dog, cat cow, horse, and other organisms.

Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the exact nucleotide sequence encoding human NOVX protein. Homologous nucleic acid sequences include those nucleic acid
5 sequences that encode conservative amino acid substitutions (see below) in SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, as well as a polypeptide possessing NOVX biological activity. Various biological activities of the NOVX proteins are described below.

A NOVX polypeptide is encoded by the open reading frame ("ORF") of a NOVX
10 nucleic acid. An ORF corresponds to a nucleotide sequence that could potentially be translated into a polypeptide. A stretch of nucleic acids comprising an ORF is uninterrupted by a stop codon. An ORF that represents the coding sequence for a full protein begins with an ATG "start" codon and terminates with one of the three "stop" codons, namely, TAA, TAG, or TGA. For the purposes of this invention, an ORF may be any part of a coding
15 sequence, with or without a start codon, a stop codon, or both. For an ORF to be considered as a good candidate for coding for a *bona fide* cellular protein, a minimum size requirement is often set, *e.g.*, a stretch of DNA that would encode a protein of 50 amino acids or more.

The nucleotide sequences determined from the cloning of the human NOVX genes allows for the generation of probes and primers designed for use in identifying and/or
20 cloning NOVX homologs in other cell types, *e.g.* from other tissues, as well as NOVX homologs from other vertebrates. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 consecutive sense strand nucleotide sequence of SEQ ID
25 NO:2*n*-1, wherein *n* is an integer between 1 and 127; or an anti-sense strand nucleotide sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127; or of a naturally occurring mutant of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127.

Probes based on the human NOVX nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various
30 embodiments, the probe has a detectable label attached, *e.g.* the label can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissues which mis-express a NOVX

protein, such as by measuring a level of a NOVX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting NOVX mRNA levels or determining whether a genomic NOVX gene has been mutated or deleted.

“A polypeptide having a biologically-active portion of a NOVX polypeptide” refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a “biologically-active portion of NOVX” can be prepared by isolating a portion of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, that encodes a polypeptide having a NOVX biological activity (the biological activities of the NOVX proteins are described below), expressing the encoded portion of NOVX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of NOVX.

NOVX Nucleic Acid and Polypeptide Variants

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequences of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, due to degeneracy of the genetic code and thus encode the same NOVX proteins as that encoded by the nucleotide sequences of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127.

In addition to the human NOVX nucleotide sequences of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the NOVX polypeptides may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the NOVX genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms “gene” and “recombinant gene” refer to nucleic acid molecules comprising an open reading frame (ORF) encoding a NOVX protein, preferably a vertebrate NOVX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the NOVX genes. Any and all such nucleotide variations and resulting amino acid polymorphisms in the NOVX polypeptides, which are

the result of natural allelic variation and that do not alter the functional activity of the NOVX polypeptides, are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding NOVX proteins from other species, and thus that have a nucleotide sequence that differs from a human SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologs of the NOVX cDNAs of the invention can be isolated based on their homology to the human NOVX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127. In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500, 750, 1000, 1500, or 2000 or more nucleotides in length. In yet another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least about 65% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.*, nucleic acids encoding NOVX proteins derived from species other than human) or other related sequences (*e.g.*, paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5 °C lower than the thermal melting point (*T_m*) for the specific sequence at a defined ionic strength and pH. The *T_m* is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at

excess, at T_m , 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30 °C for short probes, primers or oligonucleotides (*e.g.*, 10 nt to 50 nt) and at least about 60 °C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in Ausubel, *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions are hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C, followed by one or more washes in 0.2X SSC, 0.01% BSA at 50 °C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to a sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are hybridization in 6X SSC, 5X Reinhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55 °C, followed by one or more washes in 1X SSC, 0.1% SDS at 37 °C. Other conditions of moderate stringency that may be used are well-known within the art. *See, e.g.*, Ausubel, *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Krieger, 1990; GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequences of SEQ ID NO:2*n*-1, wherein *n* is an integer

between 1 and 127, or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10%
5 (wt/vol) dextran sulfate at 40 °C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50 °C. Other conditions of low stringency that may be used are well known in the art (*e.g.*, as employed for cross-species hybridizations). *See, e.g.*, Ausubel, *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND
10 EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981. *Proc Natl Acad Sci USA* 78: 6789-6792.

Conservative Mutations

In addition to naturally-occurring allelic variants of NOVX sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced
15 by mutation into the nucleotide sequences of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, thereby leading to changes in the amino acid sequences of the encoded NOVX protein, without altering the functional ability of that NOVX protein. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:2*n*, wherein *n* is an integer between 1
20 and 127. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequences of the NOVX proteins without altering their biological activity, whereas an "essential" amino acid residue is required for such biological activity. For example, amino acid residues that are conserved among the NOVX proteins of the invention are predicted to be particularly non-amenable to alteration. Amino acids for which
25 conservative substitutions can be made are well-known within the art.

Another aspect of the invention pertains to nucleic acid molecules encoding NOVX proteins that contain changes in amino acid residues that are not essential for activity. Such NOVX proteins differ in amino acid sequence from SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, yet retain biological activity. In one embodiment, the isolated
30 nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 40% homologous to the amino acid sequences of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127. Preferably, the

protein encoded by the nucleic acid molecule is at least about 60% homologous to SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127; more preferably at least about 70% homologous to SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127; still more preferably at least about 80% homologous to SEQ ID NO:2*n*, wherein *n* is an integer
5 between 1 and 127; even more preferably at least about 90% homologous to SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127; and most preferably at least about 95% homologous to SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127.

An isolated nucleic acid molecule encoding a NOVX protein homologous to the protein of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127, can be created by
10 introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced any one of SEQ ID NO:2*n*-1, wherein *n* is an integer
15 between 1 and 127, by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted, non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have
20 been defined within the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and
25 aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted non-essential amino acid residue in the NOVX protein is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a NOVX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for NOVX biological
30 activity to identify mutants that retain activity. Following mutagenesis of a nucleic acid of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, the encoded protein can be

expressed by any recombinant technology known in the art and the activity of the protein can be determined.

The relatedness of amino acid families may also be determined based on side chain interactions. Substituted amino acids may be fully conserved “strong” residues or fully conserved “weak” residues. The “strong” group of conserved amino acid residues may be any one of the following groups: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW, wherein the single letter amino acid codes are grouped by those amino acids that may be substituted for each other. Likewise, the “weak” group of conserved residues may be any one of the following: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, HFY, wherein the letters within each group represent the single letter amino acid code.

In one embodiment, a mutant NOVX protein can be assayed for (i) the ability to form protein:protein interactions with other NOVX proteins, other cell-surface proteins, or biologically-active portions thereof, (ii) complex formation between a mutant NOVX protein and a NOVX ligand; or (iii) the ability of a mutant NOVX protein to bind to an intracellular target protein or biologically-active portion thereof; (*e.g.* avidin proteins).

In yet another embodiment, a mutant NOVX protein can be assayed for the ability to regulate a specific biological function (*e.g.*, regulation of insulin release).

Interfering RNA

In one aspect of the invention, NOVX gene expression can be attenuated by RNA interference. One approach well-known in the art is short interfering RNA (siRNA) mediated gene silencing where expression products of a NOVX gene are targeted by specific double stranded NOVX derived siRNA nucleotide sequences that are complementary to at least a 19-25 nt long segment of the NOVX gene transcript, including the 5' untranslated (UT) region, the ORF, or the 3' UT region. *See, e.g.*, PCT applications WO00/44895, WO99/32619, WO01/75164, WO01/92513, WO 01/29058, WO01/89304, WO02/16620, and WO02/29858, each incorporated by reference herein in their entirety. Targeted genes can be a NOVX gene, or an upstream or downstream modulator of the NOVX gene. Nonlimiting examples of upstream or downstream modulators of a NOVX gene include, *e.g.*, a transcription factor that binds the NOVX gene promoter, a kinase or

phosphatase that interacts with a NOVX polypeptide, and polypeptides involved in a NOVX regulatory pathway.

According to the methods of the present invention, NOVX gene expression is silenced using short interfering RNA. A NOVX polynucleotide according to the invention
5 includes a siRNA polynucleotide. Such a NOVX siRNA can be obtained using a NOVX polynucleotide sequence, for example, by processing the NOVX ribopolynucleotide sequence in a cell-free system, such as but not limited to a *Drosophila* extract, or by transcription of recombinant double stranded NOVX RNA or by chemical synthesis of nucleotide sequences homologous to a NOVX sequence. *See, e.g.*, Tuschl, Zamore,
10 Lehmann, Bartel and Sharp (1999), *Genes & Dev.* 13: 3191-3197, incorporated herein by reference in its entirety. When synthesized, a typical 0.2 micromolar-scale RNA synthesis provides about 1 milligram of siRNA, which is sufficient for 1000 transfection experiments using a 24-well tissue culture plate format.

The most efficient silencing is generally observed with siRNA duplexes composed
15 of a 21-nt sense strand and a 21-nt antisense strand, paired in a manner to have a 2-nt 3' overhang. The sequence of the 2-nt 3' overhang makes an additional small contribution to the specificity of siRNA target recognition. The contribution to specificity is localized to the unpaired nucleotide adjacent to the first paired bases. In one embodiment, the nucleotides in the 3' overhang are ribonucleotides. In an alternative embodiment, the nucleotides in the
20 3' overhang are deoxyribonucleotides. Using 2'-deoxyribonucleotides in the 3' overhangs is as efficient as using ribonucleotides, but deoxyribonucleotides are often cheaper to synthesize and are most likely more nuclease resistant.

A contemplated recombinant expression vector of the invention comprises a NOVX DNA molecule cloned into an expression vector comprising operatively-linked regulatory
25 sequences flanking the NOVX sequence in a manner that allows for expression (by transcription of the DNA molecule) of both strands. An RNA molecule that is antisense to NOVX mRNA is transcribed by a first promoter (*e.g.*, a promoter sequence 3' of the cloned DNA) and an RNA molecule that is the sense strand for the NOVX mRNA is transcribed by a second promoter (*e.g.*, a promoter sequence 5' of the cloned DNA). The sense and
30 antisense strands may hybridize *in vivo* to generate siRNA constructs for silencing of the NOVX gene. Alternatively, two constructs can be utilized to create the sense and antisense strands of a siRNA construct. Finally, cloned DNA can encode a construct having secondary structure, wherein a single transcript has both the sense and complementary

antisense sequences from the target gene or genes. In an example of this embodiment, a hairpin RNAi product is homologous to all or a portion of the target gene. In another example, a hairpin RNAi product is a siRNA. The regulatory sequences flanking the NOVX sequence may be identical or may be different, such that their expression may be modulated independently, or in a temporal or spatial manner.

In a specific embodiment, siRNAs are transcribed intracellularly by cloning the NOVX gene templates into a vector containing, *e.g.*, a RNA pol III transcription unit from the smaller nuclear RNA (snRNA) U6 or the human RNase P RNA H1. One example of a vector system is the GeneSuppressorTM RNA Interference kit (commercially available from Imgenex). The U6 and H1 promoters are members of the type III class of Pol III promoters. The +1 nucleotide of the U6-like promoters is always guanosine, whereas the +1 for H1 promoters is adenosine. The termination signal for these promoters is defined by five consecutive thymidines. The transcript is typically cleaved after the second uridine. Cleavage at this position generates a 3' UU overhang in the expressed siRNA, which is similar to the 3' overhangs of synthetic siRNAs. Any sequence less than 400 nucleotides in length can be transcribed by these promoter, therefore they are ideally suited for the expression of around 21-nucleotide siRNAs in, *e.g.*, an approximately 50-nucleotide RNA stem-loop transcript.

A siRNA vector appears to have an advantage over synthetic siRNAs where long term knock-down of expression is desired. Cells transfected with a siRNA expression vector would experience steady, long-term mRNA inhibition. In contrast, cells transfected with exogenous synthetic siRNAs typically recover from mRNA suppression within seven days or ten rounds of cell division. The long-term gene silencing ability of siRNA expression vectors may provide for applications in gene therapy.

In general, siRNAs are chopped from longer dsRNA by an ATP-dependent ribonuclease called DICER. DICER is a member of the RNase III family of double-stranded RNA-specific endonucleases. The siRNAs assemble with cellular proteins into an endonuclease complex. *In vitro* studies in *Drosophila* suggest that the siRNAs/protein complex (siRNP) is then transferred to a second enzyme complex, called an RNA-induced silencing complex (RISC), which contains an endoribonuclease that is distinct from DICER. RISC uses the sequence encoded by the antisense siRNA strand to find and destroy mRNAs of complementary sequence. The siRNA thus acts as a guide, restricting the ribonuclease to cleave only mRNAs complementary to one of the two siRNA strands.

A NOVX mRNA region to be targeted by siRNA is generally selected from a desired NOVX sequence beginning 50 to 100 nt downstream of the start codon. Alternatively, 5' or 3' UTRs and regions nearby the start codon can be used but are generally avoided, as these may be richer in regulatory protein binding sites. UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP or RISC endonuclease complex. An initial BLAST homology search for the selected siRNA sequence is done against an available nucleotide sequence library to ensure that only one gene is targeted. Specificity of target recognition by siRNA duplexes indicate that a single point mutation located in the paired region of an siRNA duplex is sufficient to abolish target mRNA degradation. See, Elbashir *et al.* 2001 EMBO J. 20(23):6877-88. Hence, consideration should be taken to accommodate SNPs, polymorphisms, allelic variants or species-specific variations when targeting a desired gene.

In one embodiment, a complete NOVX siRNA experiment includes the proper negative control. A negative control siRNA generally has the same nucleotide composition as the NOVX siRNA but lack significant sequence homology to the genome. Typically, one would scramble the nucleotide sequence of the NOVX siRNA and do a homology search to make sure it lacks homology to any other gene.

Two independent NOVX siRNA duplexes can be used to knock-down a target NOVX gene. This helps to control for specificity of the silencing effect. In addition, expression of two independent genes can be simultaneously knocked down by using equal concentrations of different NOVX siRNA duplexes, *e.g.*, a NOVX siRNA and an siRNA for a regulator of a NOVX gene or polypeptide. Availability of siRNA-associating proteins is believed to be more limiting than target mRNA accessibility.

A targeted NOVX region is typically a sequence of two adenines (AA) and two thymidines (TT) divided by a spacer region of nineteen (N19) residues (*e.g.*, AA(N19)TT). A desirable spacer region has a G/C-content of approximately 30% to 70%, and more preferably of about 50%. If the sequence AA(N19)TT is not present in the target sequence, an alternative target region would be AA(N21). The sequence of the NOVX sense siRNA corresponds to (N19)TT or N21, respectively. In the latter case, conversion of the 3' end of the sense siRNA to TT can be performed if such a sequence does not naturally occur in the NOVX polynucleotide. The rationale for this sequence conversion is to generate a symmetric duplex with respect to the sequence composition of the sense and antisense 3' overhangs. Symmetric 3' overhangs may help to ensure that the siRNPs are formed with

approximately equal ratios of sense and antisense target RNA-cleaving siRNPs. *See, e.g.,* Elbashir, Lendeckel and Tuschl (2001). *Genes & Dev.* 15: 188-200, incorporated by reference herein in its entirety. The modification of the overhang of the sense sequence of the siRNA duplex is not expected to affect targeted mRNA recognition, as the antisense
5 siRNA strand guides target recognition.

Alternatively, if the NOVX target mRNA does not contain a suitable AA(N21) sequence, one may search for the sequence NA(N21). Further, the sequence of the sense strand and antisense strand may still be synthesized as 5' (N19)TT, as it is believed that the sequence of the 3'-most nucleotide of the antisense siRNA does not contribute to specificity.
10 Unlike antisense or ribozyme technology, the secondary structure of the target mRNA does not appear to have a strong effect on silencing. *See, Harborth, et al. (2001) J. Cell Science* 114: 4557-4565, incorporated by reference in its entirety.

Transfection of NOVX siRNA duplexes can be achieved using standard nucleic acid transfection methods, for example, OLIGOFECTAMINE Reagent (commercially available
15 from Invitrogen). An assay for NOVX gene silencing is generally performed approximately 2 days after transfection. No NOVX gene silencing has been observed in the absence of transfection reagent, allowing for a comparative analysis of the wild-type and silenced NOVX phenotypes. In a specific embodiment, for one well of a 24-well plate, approximately 0.84 μ g of the siRNA duplex is generally sufficient. Cells are typically
20 seeded the previous day, and are transfected at about 50% confluence. The choice of cell culture media and conditions are routine to those of skill in the art, and will vary with the choice of cell type. The efficiency of transfection may depend on the cell type, but also on the passage number and the confluency of the cells. The time and the manner of formation of siRNA-liposome complexes (*e.g.* inversion versus vortexing) are also critical. Low
25 transfection efficiencies are the most frequent cause of unsuccessful NOVX silencing. The efficiency of transfection needs to be carefully examined for each new cell line to be used. Preferred cell are derived from a mammal, more preferably from a rodent such as a rat or mouse, and most preferably from a human. Where used for therapeutic treatment, the cells are preferentially autologous, although non-autologous cell sources are also contemplated as
30 within the scope of the present invention.

For a control experiment, transfection of 0.84 μ g single-stranded sense NOVX siRNA will have no effect on NOVX silencing, and 0.84 μ g antisense siRNA has a weak silencing effect when compared to 0.84 μ g of duplex siRNAs. Control experiments again

allow for a comparative analysis of the wild-type and silenced NOVX phenotypes. To control for transfection efficiency, targeting of common proteins is typically performed, for example targeting of lamin A/C or transfection of a CMV-driven EGFP-expression plasmid (e.g. commercially available from Clontech). In the above example, a determination of the
5 fraction of lamin A/C knockdown in cells is determined the next day by such techniques as immunofluorescence, Western blot, Northern blot or other similar assays for protein expression or gene expression. Lamin A/C monoclonal antibodies may be obtained from Santa Cruz Biotechnology.

Depending on the abundance and the half life (or turnover) of the targeted NOVX
10 polynucleotide in a cell, a knock-down phenotype may become apparent after 1 to 3 days, or even later. In cases where no NOVX knock-down phenotype is observed, depletion of the NOVX polynucleotide may be observed by immunofluorescence or Western blotting. If the NOVX polynucleotide is still abundant after 3 days, cells need to be split and transferred to a fresh 24-well plate for re-transfection. If no knock-down of the targeted protein is
15 observed, it may be desirable to analyze whether the target mRNA (NOVX or a NOVX upstream or downstream gene) was effectively destroyed by the transfected siRNA duplex. Two days after transfection, total RNA is prepared, reverse transcribed using a target-specific primer, and PCR-amplified with a primer pair covering at least one exon-exon junction in order to control for amplification of pre-mRNAs. RT/PCR of a non-targeted
20 mRNA is also needed as control. Effective depletion of the mRNA yet undetectable reduction of target protein may indicate that a large reservoir of stable NOVX protein may exist in the cell. Multiple transfection in sufficiently long intervals may be necessary until the target protein is finally depleted to a point where a phenotype may become apparent. If multiple transfection steps are required, cells are split 2 to 3 days after transfection. The
25 cells may be transfected immediately after splitting.

An inventive therapeutic method of the invention contemplates administering a NOVX siRNA construct as therapy to compensate for increased or aberrant NOVX expression or activity. The NOVX ribopolynucleotide is obtained and processed into siRNA fragments, or a NOVX siRNA is synthesized, as described above. The NOVX
30 siRNA is administered to cells or tissues using known nucleic acid transfection techniques, as described above. A NOVX siRNA specific for a NOVX gene will decrease or knockdown NOVX transcription products, which will lead to reduced NOVX polypeptide production, resulting in reduced NOVX polypeptide activity in the cells or tissues.

The present invention also encompasses a method of treating a disease or condition associated with the presence of a NOVX protein in an individual comprising administering to the individual an RNAi construct that targets the mRNA of the protein (the mRNA that encodes the protein) for degradation. A specific RNAi construct includes a siRNA or a
5 double stranded gene transcript that is processed into siRNAs. Upon treatment, the target protein is not produced or is not produced to the extent it would be in the absence of the treatment.

Where the NOVX gene function is not correlated with a known phenotype, a control sample of cells or tissues from healthy individuals provides a reference standard for
10 determining NOVX expression levels. Expression levels are detected using the assays described, *e.g.*, RT-PCR, Northern blotting, Western blotting, ELISA, and the like. A subject sample of cells or tissues is taken from a mammal, preferably a human subject, suffering from a disease state. The NOVX ribopolynucleotide is used to produce siRNA constructs, that are specific for the NOVX gene product. These cells or tissues are treated
15 by administering NOVX siRNA's to the cells or tissues by methods described for the transfection of nucleic acids into a cell or tissue, and a change in NOVX polypeptide or polynucleotide expression is observed in the subject sample relative to the control sample, using the assays described. This NOVX gene knockdown approach provides a rapid method for determination of a NOVX minus (NOVX⁻) phenotype in the treated subject
20 sample. The NOVX⁻ phenotype observed in the treated subject sample thus serves as a marker for monitoring the course of a disease state during treatment.

In specific embodiments, a NOVX siRNA is used in therapy. Methods for the generation and use of a NOVX siRNA are known to those skilled in the art. Example techniques are provided below.

25 **Production of RNAs**

Sense RNA (ssRNA) and antisense RNA (asRNA) of NOVX are produced using known methods such as transcription in RNA expression vectors. In the initial experiments, the sense and antisense RNA are about 500 bases in length each. The produced ssRNA and asRNA (0.5 μ M) in 10 mM Tris-HCl (pH 7.5) with 20 mM NaCl were heated to 95° C for 1
30 min then cooled and annealed at room temperature for 12 to 16 h. The RNAs are precipitated and resuspended in lysis buffer (below). To monitor annealing, RNAs are electrophoresed in a 2% agarose gel in TBE buffer and stained with ethidium bromide. See,

e.g., Sambrook *et al.*, Molecular Cloning. Cold Spring Harbor Laboratory Press, Plainview, N.Y. (1989).

Lysate Preparation

5 Untreated rabbit reticulocyte lysate (Ambion) are assembled according to the manufacturer's directions. dsRNA is incubated in the lysate at 30° C for 10 min prior to the addition of mRNAs. Then NOVX mRNAs are added and the incubation continued for an additional 60 min. The molar ratio of double stranded RNA and mRNA is about 200:1. The NOVX mRNA is radiolabeled (using known techniques) and its stability is monitored by gel electrophoresis.

10 In a parallel experiment made with the same conditions, the double stranded RNA is internally radiolabeled with a ³²P-ATP. Reactions are stopped by the addition of 2 X proteinase K buffer and deproteinized as described previously (Tuschl *et al.*, Genes Dev., 13:3191-3197 (1999)). Products are analyzed by electrophoresis in 15% or 18% polyacrylamide sequencing gels using appropriate RNA standards. By monitoring the gels
15 for radioactivity, the natural production of 10 to 25 nt RNAs from the double stranded RNA can be determined.

 The band of double stranded RNA, about 21-23 bps, is eluded. The efficacy of these 21-23 mers for suppressing NOVX transcription is assayed in vitro using the same rabbit reticulocyte assay described above using 50 nanomolar of double stranded 21-23 mer for
20 each assay. The sequence of these 21-23 mers is then determined using standard nucleic acid sequencing techniques.

RNA Preparation

 21 nt RNAs, based on the sequence determined above, are chemically synthesized using Expedite RNA phosphoramidites and thymidine phosphoramidite (Proligo, Germany).
25 Synthetic oligonucleotides are deprotected and gel-purified (Elbashir, Lendeckel, & Tuschl, Genes & Dev. 15, 188-200 (2001)), followed by Sep-Pak C18 cartridge (Waters, Milford, Mass., USA) purification (Tuschl, *et al.*, Biochemistry, 32:11658-11668 (1993)).

 These RNAs (20 μM) single strands are incubated in annealing buffer (100 mM potassium acetate, 30 mM HEPES-KOH at pH 7.4, 2 mM magnesium acetate) for 1 min at
30 90° C followed by 1 h at 37° C.

Cell Culture

A cell culture known in the art to regularly express NOVX is propagated using standard conditions. 24 hours before transfection, at approx. 80% confluency, the cells are trypsinized and diluted 1:5 with fresh medium without antibiotics (1-3 X 10⁵ cells/ml) and transferred to 24-well plates (500 µl/well). Transfection is performed using a commercially available lipofection kit and NOVX expression is monitored using standard techniques with positive and negative control. A positive control is cells that naturally express NOVX while a negative control is cells that do not express NOVX. Base-paired 21 and 22 nt siRNAs with overhanging 3' ends mediate efficient sequence-specific mRNA degradation in lysates and in cell culture. Different concentrations of siRNAs are used. An efficient concentration for suppression in vitro in mammalian culture is between 25 nM to 100 nM final concentration. This indicates that siRNAs are effective at concentrations that are several orders of magnitude below the concentrations applied in conventional antisense or ribozyme gene targeting experiments.

The above method provides a way both for the deduction of NOVX siRNA sequence and the use of such siRNA for in vitro suppression. In vivo suppression may be performed using the same siRNA using well known in vivo transfection or gene therapy transfection techniques.

Antisense Nucleic Acids

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein (e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence). In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire NOVX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a NOVX protein of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127, or antisense

nucleic acids complementary to a NOVX nucleic acid sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding a NOVX protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding the NOVX protein. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding the NOVX protein disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of NOVX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of NOVX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of NOVX mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally-occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids (*e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used).

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-carboxymethylaminomethyl-2-thiouridine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 5-methoxyuracil, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine,

5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, 2-thiouracil, 4-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 5-methyluracil, 5 uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a NOVX protein to thereby inhibit expression of the protein (*e.g.*, by inhibiting transcription and/or translation). The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface (*e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens). The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient nucleic acid molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other. *See, e.g., Gaultier, et al., 1987. Nucl. Acids Res. 15: 6625-6641.* The antisense nucleic acid molecule can also comprise a

2'-o-methylribonucleotide (See, e.g., Inoue, *et al.* 1987. *Nucl. Acids Res.* **15**: 6131-6148) or a chimeric RNA-DNA analogue (See, e.g., Inoue, *et al.*, 1987. *FEBS Lett.* **215**: 327-330).

Ribozymes and PNA Moieties

Nucleic acid modifications include, by way of non-limiting example, modified
5 bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In one embodiment, an antisense nucleic acid of the invention is a ribozyme.
10 Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes as described in Haselhoff and Gerlach 1988. *Nature* 334: 585-591) can be used to catalytically cleave NOVX mRNA transcripts to thereby inhibit translation of NOVX mRNA. A ribozyme
15 having specificity for a NOVX-encoding nucleic acid can be designed based upon the nucleotide sequence of a NOVX cDNA disclosed herein (*i.e.*, SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a NOVX-encoding mRNA. See,
20 e.g., U.S. Patent 4,987,071 to Cech, *et al.* and U.S. Patent 5,116,742 to Cech, *et al.* NOVX mRNA can also be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, NOVX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the NOVX nucleic acid (e.g., the
25 NOVX promoter and/or enhancers) to form triple helical structures that prevent transcription of the NOVX gene in target cells. See, e.g., Helene, 1991. *Anticancer Drug Des.* 6: 569-84; Helene, *et al.* 1992. *Ann. N.Y. Acad. Sci.* 660: 27-36; Maher, 1992. *Bioassays* 14: 807-15.

In various embodiments, the NOVX nucleic acids can be modified at the base
30 moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic

acids can be modified to generate peptide nucleic acids. See, e.g., Hyrup, *et al.*, 1996. *Bioorg Med Chem* 4: 5-23. As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics (e.g., DNA mimics) in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleotide
5 bases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomer can be performed using standard solid phase peptide synthesis protocols as described in Hyrup, *et al.*, 1996, *supra*; Perry-O'Keefe, *et al.*, 1996, *Proc. Natl. Acad. Sci. USA* 93: 14670-14675.

10 PNAs of NOVX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of NOVX can also be used, for example, in the analysis of single base pair mutations in a gene (e.g., PNA directed PCR clamping; as artificial
15 restriction enzymes when used in combination with other enzymes, e.g., S_I nucleases (See, Hyrup, *et al.*, 1996, *supra*); or as probes or primers for DNA sequence and hybridization (See, Hyrup, *et al.*, 1996, *supra*; Perry-O'Keefe, *et al.*, 1996, *supra*).

In another embodiment, PNAs of NOVX can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the
20 formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of NOVX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes (e.g., RNase H and DNA polymerases) to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity.
25 PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleotide bases, and orientation (see, Hyrup, *et al.*, 1996, *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup, *et al.*, 1996, *supra* and Finn, *et al.*, 1996, *Nucl Acids Res* 24: 3357-3363. For example, a DNA chain can be synthesized on a solid support using standard
30 phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA. See, e.g., Mag, *et al.*, 1989. *Nucl Acid Res* 17: 5973-5988.

PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment. *See, e.g., Finn, et al., 1996, supra.* Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. *See, e.g., Petersen, et al., 1975. Bioorg. Med. Chem. Lett. 5: 1119-1124.*

5 In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g., for targeting host cell receptors in vivo*), or agents facilitating transport across the cell membrane (*see, e.g., Letsinger, et al., 1989. Proc. Natl. Acad. Sci. U.S.A. 86: 6553-6556; Lemaitre, et al., 1987. Proc. Natl. Acad. Sci. 84: 648-652; PCT Publication No. WO88/09810*) or the blood-brain barrier (*see, e.g., PCT Publication No. WO 89/10134*). In
10 addition, oligonucleotides can be modified with hybridization triggered cleavage agents (*see, e.g., Krol, et al., 1988. BioTechniques 6:958-976*) or intercalating agents (*see, e.g., Zon, 1988. Pharm. Res. 5: 539-549*). To this end, the oligonucleotide may be conjugated to another molecule, *e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, and the like.*

15 **NOVX Polypeptides**

A polypeptide according to the invention includes a polypeptide including the amino acid sequence of NOVX polypeptides whose sequences are provided in any one of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127. The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residues
20 shown in any one of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127, while still encoding a protein that maintains its NOVX activities and physiological functions, or a functional fragment thereof.

In general, a NOVX variant that preserves NOVX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other
25 amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above.

30 One aspect of the invention pertains to isolated NOVX proteins, and biologically-active portions thereof, or derivatives, fragments, analogs or homologs thereof.

Also provided are polypeptide fragments suitable for use as immunogens to raise anti-NOVX antibodies. In one embodiment, native NOVX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, NOVX proteins are produced by
5 recombinant DNA techniques. Alternative to recombinant expression, a NOVX protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" polypeptide or protein or biologically-active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the NOVX protein is derived, or substantially free from
10 chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of NOVX proteins in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly-produced. In one embodiment, the language "substantially free of cellular material" includes preparations of NOVX proteins having less than about 30% (by dry
15 weight) of non-NOVX proteins (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-NOVX proteins, still more preferably less than about 10% of non-NOVX proteins, and most preferably less than about 5% of non-NOVX proteins. When the NOVX protein or biologically-active portion thereof is recombinantly-produced, it is also preferably substantially free of culture medium, *i.e.*,
20 culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the NOVX protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of NOVX proteins in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the
25 language "substantially free of chemical precursors or other chemicals" includes preparations of NOVX proteins having less than about 30% (by dry weight) of chemical precursors or non-NOVX chemicals, more preferably less than about 20% chemical precursors or non-NOVX chemicals, still more preferably less than about 10% chemical precursors or non-NOVX chemicals, and most preferably less than about 5% chemical
30 precursors or non-NOVX chemicals.

Biologically-active portions of NOVX proteins include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequences of the

- NOVX proteins (*e.g.*, the amino acid sequence of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127) that include fewer amino acids than the full-length NOVX proteins, and exhibit at least one activity of a NOVX protein. Typically, biologically-active portions comprise a domain or motif with at least one activity of the NOVX protein. A
- 5 biologically-active portion of a NOVX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acid residues in length.

Moreover, other biologically-active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native NOVX protein.

- 10 In an embodiment, the NOVX protein has an amino acid sequence of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127. In other embodiments, the NOVX protein is substantially homologous to SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127, and retains the functional activity of the protein of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127, yet differs in amino acid sequence due to natural allelic variation
- 15 or mutagenesis, as described in detail, below. Accordingly, in another embodiment, the NOVX protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127, and retains the functional activity of the NOVX proteins of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127.

20 **Determining Homology Between Two or More Sequences**

- To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or
- 25 nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

- 30 The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs

known in the art, such as GAP software provided in the GCG program package. *See*, Needleman and Wunsch, 1970. *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid
5 sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127.

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of
10 comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of
15 comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to
20 a reference sequence over a comparison region.

Chimeric and Fusion Proteins

The invention also provides NOVX chimeric or fusion proteins. As used herein, a NOVX "chimeric protein" or "fusion protein" comprises a NOVX polypeptide operatively-linked to a non-NOVX polypeptide. An "NOVX polypeptide" refers to a
25 polypeptide having an amino acid sequence corresponding to a NOVX protein of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127, whereas a "non-NOVX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the NOVX protein, *e.g.*, a protein that is different from the NOVX protein and that is derived from the same or a different organism. Within a NOVX
30 fusion protein the NOVX polypeptide can correspond to all or a portion of a NOVX protein. In one embodiment, a NOVX fusion protein comprises at least one biologically-active portion of a NOVX protein. In another embodiment, a NOVX fusion protein comprises at

least two biologically-active portions of a NOVX protein. In yet another embodiment, a NOVX fusion protein comprises at least three biologically-active portions of a NOVX protein. Within the fusion protein, the term "operatively-linked" is intended to indicate that the NOVX polypeptide and the non-NOVX polypeptide are fused in-frame with one
5 another. The non-NOVX polypeptide can be fused to the N-terminus or C-terminus of the NOVX polypeptide.

In one embodiment, the fusion protein is a GST-NOVX fusion protein in which the NOVX sequences are fused to the C-terminus of the GST (glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant NOVX
10 polypeptides.

In another embodiment, the fusion protein is a NOVX protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of NOVX can be increased through use of a heterologous signal sequence.

15 In yet another embodiment, the fusion protein is a NOVX-immunoglobulin fusion protein in which the NOVX sequences are fused to sequences derived from a member of the immunoglobulin protein family. The NOVX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a NOVX ligand and a NOVX protein on the
20 surface of a cell, to thereby suppress NOVX-mediated signal transduction *in vivo*. The NOVX-immunoglobulin fusion proteins can be used to affect the bioavailability of a NOVX cognate ligand. Inhibition of the NOVX ligand/NOVX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the
25 NOVX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-NOVX antibodies in a subject, to purify NOVX ligands, and in screening assays to identify molecules that inhibit the interaction of NOVX with a NOVX ligand.

A NOVX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different
30 polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as

appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary
5 overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (*see, e.g.,* Ausubel, *et al.* (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.,* a GST polypeptide). A NOVX-encoding nucleic acid can be cloned into such an expression
10 vector such that the fusion moiety is linked in-frame to the NOVX protein.

NOVX Agonists and Antagonists

The invention also pertains to variants of the NOVX proteins that function as either NOVX agonists (*i.e.,* mimetics) or as NOVX antagonists. Variants of the NOVX protein can be generated by mutagenesis (*e.g.,* discrete point mutation or truncation of the NOVX
15 protein). An agonist of the NOVX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the NOVX protein. An antagonist of the NOVX protein can inhibit one or more of the activities of the naturally occurring form of the NOVX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the NOVX
20 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the NOVX proteins.

25 Variants of the NOVX proteins that function as either NOVX agonists (*i.e.,* mimetics) or as NOVX antagonists can be identified by screening combinatorial libraries of mutants (*e.g.,* truncation mutants) of the NOVX proteins for NOVX protein agonist or antagonist activity. In one embodiment, a variegated library of NOVX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene
30 library. A variegated library of NOVX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential NOVX sequences is expressible as individual polypeptides, or

alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of NOVX sequences therein. There are a variety of methods which can be used to produce libraries of potential NOVX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential NOVX sequences. Methods for synthesizing degenerate oligonucleotides are well-known within the art. *See, e.g.*, Narang, 1983. *Tetrahedron* 39: 3; Itakura, *et al.*, 1984. *Annu. Rev. Biochem.* 53: 323; Itakura, *et al.*, 1984. *Science* 198: 1056; Ike, *et al.*, 1983. *Nucl. Acids Res.* 11: 477.

Polypeptide Libraries

In addition, libraries of fragments of the NOVX protein coding sequences can be used to generate a variegated population of NOVX fragments for screening and subsequent selection of variants of a NOVX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a NOVX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double-stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S_1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, expression libraries can be derived which encodes N-terminal and internal fragments of various sizes of the NOVX proteins.

Various techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of NOVX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the

frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify NOVX variants. See, e.g., Arkin and Youvan, 1992, *Proc. Natl. Acad. Sci. USA* 89: 7811-7815; Delgrave, *et al.*, 1993. *Protein Engineering* 6:327-331.

Anti-NOVX Antibodies

5 Included in the invention are antibodies to NOVX proteins, or fragments of NOVX proteins. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single
10 chain, F_{ab} , F_{ab}' and $F_{(ab)2}$ fragments, and an F_{ab} expression library. In general, antibody molecules obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG₁, IgG₂, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to
15 antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated protein of the invention intended to serve as an antigen, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and
20 monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as an amino acid sequence of SEQ ID NO:2*n*, wherein *n* is an integer between 1 and 127, and encompasses an epitope thereof such that an
25 antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly
30 these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of NOVX that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human NOVX protein sequence will indicate which regions of a NOVX polypeptide are particularly hydrophilic and, therefore,
5 are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. *See, e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78:
10 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each incorporated herein by reference in their entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

The term "epitope" includes any protein determinant capable of specific binding to
15 an immunoglobulin or T-cell receptor. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics. A NOVX polypeptide or a fragment thereof comprises at least one antigenic epitope. An anti-NOVX antibody of the present invention is said to specifically bind to
20 antigen NOVX when the equilibrium binding constant (K_D) is $\leq 1 \mu\text{M}$, preferably $\leq 100 \text{ nM}$, more preferably $\leq 10 \text{ nM}$, and most preferably $\leq 100 \text{ pM}$ to about 1 pM , as measured by assays such as radioligand binding assays or similar assays known to those skilled in the art.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that
25 immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (*see, for example*,
Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor
30 Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

Polyclonal Antibodies

For the production of polyclonal antibodies, various suitable host animals (*e.g.*, rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (*e.g.*, aluminum hydroxide), surface active substances (*e.g.*, lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, *etc.*), adjuvants usable in humans such as Bacille Calmette-Guerin and *Corynebacterium parvum*, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (*The Scientist*, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

Monoclonal Antibodies

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique

heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MABs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

5 Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be
10 immunized in vitro.

 The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell
15 line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that
20 preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

25 Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas,
30 Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, *J. Immunol.*,

133:3001 (1984); Brodeur *et al.*, Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the
5 binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). It is an
10 objective, especially important in therapeutic applications of monoclonal antibodies, to identify antibodies having a high degree of specificity and a high binding affinity for the target antigen.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods (Goding, 1986). Suitable
15 culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification
20 procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (*e.g.*, by
25 using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce
30 immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the

homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or
5 can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

Humanized Antibodies

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for
10 administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a
15 non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones *et al.*, Nature, 321:522-525 (1986); Riechmann *et al.*, Nature, 332:323-327 (1988); Verhoeven *et al.*, Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human
20 immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human
25 immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones *et al.*, 1986; Riechmann *et al.*, 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

Human Antibodies

Fully human antibodies essentially relate to antibody molecules in which the entire sequence of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

- 5 Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, *et al.*, 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be
- 10 produced by using human hybridomas (see Cote, *et al.*, 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

- In addition, human antibodies can also be produced using additional techniques,
- 15 including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks *et al.*, J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, *e.g.*, mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in
- 20 humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks *et al.* (Bio/Technology 10, 779-783 (1992)); Lonberg *et al.* (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild *et al.* (Nature Biotechnology 14, 845-51 (1996)); Neuberger
- 25 (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

- Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT
- 30 publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's

genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the XenomouseTM as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

F_{ab} Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression
5 libraries (see *e.g.*, Huse, *et al.*, 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an F_{(ab')₂} fragment produced by pepsin digestion of an antibody
10 molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an F_{(ab')₂} fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies
15 that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the
20 recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the
25 correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen
30 combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part

of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are
5 co-transfected into a suitable host organism. For further details of generating bispecific antibodies *see*, for example, Suresh *et al.*, Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at
10 least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (*e.g.* tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (*e.g.* alanine or
15 threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (*e.g.* F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies
20 can be prepared using chemical linkage. Brennan *et al.*, Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives.
25 One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from *E. coli* and chemically
30 coupled to form bispecific antibodies. Shalaby *et al.*, J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical

coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

5 Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny *et al.*, J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were
10 reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger *et al.*, Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a
15 light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, J. Immunol. 152:5368 (1994).
20

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt *et al.*, J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm
25 of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (*e.g.* CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a
30 particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or

TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention.

- 5 Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking
- 10 agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

Effector Function Engineering

- 15 It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased
- 20 complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual
- 25 Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, Anti-Cancer Drug Design, 3: 219-230 (1989).

Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (*e.g.*, an

enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used
5 include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolacca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the
10 tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y , and ^{186}Re .

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl
15 adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaredehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as
20 described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such
25 streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.*, avidin) that is in turn conjugated to a cytotoxic agent.

Immunoliposomes

30 The antibodies disclosed herein can also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as

described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse-phase evaporation
5 method with a lipid composition comprising phosphatidylcholine, cholesterol, and
PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through
filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of
the antibody of the present invention can be conjugated to the liposomes as described in
Martin *et al.*, J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A
10 chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome.
See Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

Diagnostic Applications of Antibodies Directed Against the Proteins of the Invention

In one embodiment, methods for the screening of antibodies that possess the desired
15 specificity include, but are not limited to, enzyme linked immunosorbent assay (ELISA) and
other immunologically mediated techniques known within the art. In a specific
embodiment, selection of antibodies that are specific to a particular domain of an NOVX
protein is facilitated by generation of hybridomas that bind to the fragment of an NOVX
protein possessing such a domain. Thus, antibodies that are specific for a desired domain
20 within an NOVX protein, or derivatives, fragments, analogs or homologs thereof, are also
provided herein.

Antibodies directed against a NOVX protein of the invention may be used in
methods known within the art relating to the localization and/or quantitation of a NOVX
protein (*e.g.*, for use in measuring levels of the NOVX protein within appropriate
25 physiological samples, for use in diagnostic methods, for use in imaging the protein, and the
like). In a given embodiment, antibodies specific to a NOVX protein, or derivative,
fragment, analog or homolog thereof, that contain the antibody derived antigen binding
domain, are utilized as pharmacologically active compounds (referred to hereinafter as
"Therapeutics").

30 An antibody specific for a NOVX protein of the invention (*e.g.*, a monoclonal
antibody or a polyclonal antibody) can be used to isolate a NOVX polypeptide by standard

- techniques, such as immunoaffinity, chromatography or immunoprecipitation. An antibody to a NOVX polypeptide can facilitate the purification of a natural NOVX antigen from cells, or of a recombinantly produced NOVX antigen expressed in host cells. Moreover, such an anti-NOVX antibody can be used to detect the antigenic NOVX protein (*e.g.*, in a cellular
- 5 lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the antigenic NOVX protein. Antibodies directed against a NOVX protein can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance.
- 10 Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials
- 15 include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

20 Antibody Therapeutics

- Antibodies of the invention, including polyclonal, monoclonal, humanized and fully human antibodies, may be used as therapeutic agents. Such agents will generally be employed to treat or prevent a disease or pathology in a subject. An antibody preparation, preferably one having high specificity and high affinity for its target antigen, is administered to the
- 25 subject and will generally have an effect due to its binding with the target. Such an effect may be one of two kinds, depending on the specific nature of the interaction between the given antibody molecule and the target antigen in question. In the first instance, administration of the antibody may abrogate or inhibit the binding of the target with an endogenous ligand to which it naturally binds. In this case, the antibody binds to the target
- 30 and masks a binding site of the naturally occurring ligand, wherein the ligand serves as an effector molecule. Thus the receptor mediates a signal transduction pathway for which ligand is responsible.

Alternatively, the effect may be one in which the antibody elicits a physiological result by virtue of binding to an effector binding site on the target molecule. In this case the target, a receptor having an endogenous ligand which may be absent or defective in the disease or pathology, binds the antibody as a surrogate effector ligand, initiating a
5 receptor-based signal transduction event by the receptor.

A therapeutically effective amount of an antibody of the invention relates generally to the amount needed to achieve a therapeutic objective. As noted above, this may be a binding interaction between the antibody and its target antigen that, in certain cases, interferes with the functioning of the target, and in other cases, promotes a physiological
10 response. The amount required to be administered will furthermore depend on the binding affinity of the antibody for its specific antigen, and will also depend on the rate at which an administered antibody is depleted from the free volume other subject to which it is administered. Common ranges for therapeutically effective dosing of an antibody or antibody fragment of the invention may be, by way of nonlimiting example, from about 0.1
15 mg/kg body weight to about 50 mg/kg body weight. Common dosing frequencies may range, for example, from twice daily to once a week.

Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a protein of the invention, as well as other molecules identified by the screening assays disclosed herein, can be administered for the treatment of
20 various disorders in the form of pharmaceutical compositions. Principles and considerations involved in preparing such compositions, as well as guidance in the choice of components are provided, for example, in Remington: The Science And Practice Of Pharmacy 19th ed. (Alfonso R. Gennaro, *et al.*, editors) Mack Pub. Co., Easton, Pa.: 1995; Drug Absorption Enhancement: Concepts, Possibilities, Limitations, And Trends, Harwood
25 Academic Publishers, Langhorne, Pa., 1994; and Peptide And Protein Drug Delivery (Advances In Parenteral Sciences, Vol. 4), 1991, M. Dekker, New York.

If the antigenic protein is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the
30 smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody,

peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. *See, e.g.,* Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993). The formulation herein can also contain more than one active compound as necessary for
5 the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition can comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

10 The active ingredients can also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in
15 macroemulsions.

The formulations to be used for in vivo administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations can be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic
20 polymers containing the antibody, which matrices are in the form of shaped articles, e.g., films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic
25 acid copolymers such as the LUPRON DEPOTTM (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods.

ELISA Assay

An agent for detecting an analyte protein is an antibody capable of binding to an analyte protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*,
5 F_{ab} or F_{(ab)2}) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently-labeled
10 secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently-labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. Included within the usage of the term "biological sample", therefore, is blood and a fraction or component of blood including blood serum,
15 blood plasma, or lymph. That is, the detection method of the invention can be used to detect an analyte mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of an analyte mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of an analyte protein include enzyme linked immunosorbent assays (ELISAs), Western blots,
20 immunoprecipitations, and immunofluorescence. *In vitro* techniques for detection of an analyte genomic DNA include Southern hybridizations. Procedures for conducting immunoassays are described, for example in "ELISA: Theory and Practice: Methods in Molecular Biology", Vol. 42, J. R. Crowther (Ed.) Human Press, Totowa, NJ, 1995; "Immunoassay", E. Diamandis and T. Christopoulos, Academic Press, Inc., San Diego, CA,
25 1996; and "Practice and Theory of Enzyme Immunoassays", P. Tijssen, Elsevier Science Publishers, Amsterdam, 1985. Furthermore, *in vivo* techniques for detection of an analyte protein include introducing into a subject a labeled anti-analyte protein antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

NOVX Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a NOVX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule

5 capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors

10 having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively-linked. Such vectors are referred to herein as "expression vectors". In general,

15 expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve

20 equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively-linked to the

25 nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably-linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell).

30 The term "regulatory sequence" is intended to includes promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY:

METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled
5 in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, *etc.* The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, NOVX proteins, mutant forms of NOVX proteins, fusion proteins,
10 *etc.*).

The recombinant expression vectors of the invention can be designed for expression of NOVX proteins in prokaryotic or eukaryotic cells. For example, NOVX proteins can be expressed in bacterial cells such as *Escherichia coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further
15 in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *Escherichia coli*
20 with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (i) to increase expression of recombinant protein; (ii) to increase the solubility of the recombinant protein; and (iii) to aid in the
25 purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical
30 fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia,

Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION
5 TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. *See, e.g.*, Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN
10 ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (*see, e.g.*, Wada, *et al.*, 1992. *Nucl. Acids Res.* 20: 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

15 In another embodiment, the NOVX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *Saccharomyces cerevisiae* include pYepSec1 (Baldari, *et al.*, 1987. *EMBO J.* 6: 229-234), pMFa (Kurjan and Herskowitz, 1982. *Cell* 30: 933-943), pJRY88 (Schultz *et al.*, 1987. *Gene* 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (InVitrogen Corp, San Diego, Calif.).

20 Alternatively, NOVX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, SF9 cells) include the pAc series (Smith, *et al.*, 1983. *Mol. Cell. Biol.* 3: 2156-2165) and the pVL series (Lucklow and Summers, 1989. *Virology* 170: 31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in
25 mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, 1987. *Nature* 329: 840) and pMT2PC (Kaufman, *et al.*, 1987. *EMBO J.* 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus, and simian virus
30 40. For other suitable expression systems for both prokaryotic and eukaryotic cells *see, e.g.*, Chapters 16 and 17 of Sambrook, *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL.

2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert, *et al.*, 1987. *Genes Dev.* 1: 268-277), lymphoid-specific promoters (Calame and Eaton, 1988. *Adv. Immunol.* 43: 235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989. *EMBO J.* 8: 729-733) and immunoglobulins (Banerji, *et al.*, 1983. *Cell* 33: 729-740; Queen and Baltimore, 1983. *Cell* 33: 741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddell, 1989. *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), pancreas-specific promoters (Edlund, *et al.*, 1985. *Science* 230: 912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, e.g., the murine hox promoters (Kessel and Gruss, 1990. *Science* 249: 374-379) and the α -fetoprotein promoter (Campes and Tilghman, 1989. *Genes Dev.* 3: 537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively-linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to NOVX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see, e.g., Weintraub, *et al.*, "Antisense RNA as a molecular tool for genetic analysis," *Reviews-Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of
5 such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, NOVX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells
10 (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing
15 foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989),
20 and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is
25 generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding NOVX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*,
30 cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) NOVX protein. Accordingly, the invention further provides methods for producing NOVX protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a
5 recombinant expression vector encoding NOVX protein has been introduced) in a suitable medium such that NOVX protein is produced. In another embodiment, the method further comprises isolating NOVX protein from the medium or the host cell.

Transgenic NOVX Animals

The host cells of the invention can also be used to produce non-human transgenic
10 animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which NOVX protein-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous NOVX sequences have been introduced into their genome or homologous recombinant animals in which endogenous NOVX sequences have been altered. Such
15 animals are useful for studying the function and/or activity of NOVX protein and for identifying and/or evaluating modulators of NOVX protein activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows,
20 goats, chickens, amphibians, *etc.* A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a
25 mouse, in which an endogenous NOVX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing NOVX-encoding
30 nucleic acid into the male pronuclei of a fertilized oocyte (*e.g.*, by microinjection, retroviral infection) and allowing the oocyte to develop in a pseudopregnant female foster animal. The human NOVX cDNA sequences, *i.e.*, any one of SEQ ID NO:2*n*-1, wherein *n* is an

integer between 1 and 127, can be introduced as a transgene into the genome of a non-human animal. Alternatively, a non-human homolog of the human NOVX gene, such as a mouse NOVX gene, can be isolated based on hybridization to the human NOVX cDNA (described further *supra*) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably-linked to the NOVX transgene to direct expression of NOVX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan, 1986. In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the NOVX transgene in its genome and/or expression of NOVX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene-encoding NOVX protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a NOVX gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the NOVX gene. The NOVX gene can be a human gene (*e.g.*, the cDNA of any one of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127), but more preferably, is a non-human homolog of a human NOVX gene. For example, a mouse homolog of human NOVX gene of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, can be used to construct a homologous recombination vector suitable for altering an endogenous NOVX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous NOVX gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous NOVX gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous NOVX protein). In the homologous recombination

vector, the altered portion of the NOVX gene is flanked at its 5'- and 3'-termini by additional nucleic acid of the NOVX gene to allow for homologous recombination to occur between the exogenous NOVX gene carried by the vector and an endogenous NOVX gene in an embryonic stem cell. The additional flanking NOVX nucleic acid is of sufficient
5 length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5'- and 3'-termini) are included in the vector. *See, e.g.,* Thomas, *et al.*, 1987. *Cell* 51: 503 for a description of homologous recombination vectors. The vector is then introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced NOVX gene has
10 homologously-recombined with the endogenous NOVX gene are selected. *See, e.g.,* Li, *et al.*, 1992. *Cell* 69: 915.

The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras. *See, e.g.,* Bradley, 1987. In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp.
15 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously-recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously-recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors
20 and homologous recombinant animals are described further in Bradley, 1991. *Curr. Opin. Biotechnol.* 2: 823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example
25 of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, *See, e.g.,* Lakso, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae*. *See, O'Gorman, et al.*, 1991. *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals
30 containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic

animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, *et al.*, 1997. *Nature* 385: 810-813. In brief, a cell (*e.g.*, a somatic cell) from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell (*e.g.*, the somatic cell) is isolated.

Pharmaceutical Compositions

The NOVX nucleic acid molecules, NOVX proteins, and anti-NOVX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include

parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (*i.e.*, topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid (EDTA); buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide.

10 The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

15
20
25
30

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a NOVX protein or anti-NOVX antibody) in the required amount in an appropriate

solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the
5 preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral
10 therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The
15 tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring
20 agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For
25 transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are
30 formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will
5 protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be
10 obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

15 It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical
20 carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

The nucleic acid molecules of the invention can be inserted into vectors and used as
25 gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see, e.g.*, U.S. Patent No. 5,328,470) or by stereotactic injection (*see, e.g.*, Chen, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which
30 the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the

pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

5 **Screening and Detection Methods**

The isolated nucleic acid molecules of the invention can be used to express NOVX protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect NOVX mRNA (*e.g.*, in a biological sample) or a genetic lesion in a NOVX gene, and to modulate NOVX activity, as described further, below. In addition, the NOVX proteins can be used to screen drugs or compounds that modulate the NOVX protein activity or expression as well as to treat disorders characterized by insufficient or excessive production of NOVX protein or production of NOVX protein forms that have decreased or aberrant activity compared to NOVX wild-type protein (*e.g.*; diabetes (regulates insulin release); obesity (binds and transport lipids); metabolic disturbances associated with obesity, the metabolic syndrome X, as well as anorexia and wasting disorders associated with chronic diseases and various cancers, and infectious disease (possesses anti-microbial activity) and the various dyslipidemias. In addition, the anti-NOVX antibodies of the invention can be used to detect and isolate NOVX proteins and modulate NOVX activity. In yet a further aspect, the invention can be used in methods to influence appetite, absorption of nutrients and the disposition of metabolic substrates in both a positive and negative fashion.

The invention further pertains to novel agents identified by the screening assays described herein and uses thereof for treatments as described, *supra*.

Screening Assays

25 The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) that bind to NOVX proteins or have a stimulatory or inhibitory effect on, *e.g.*, NOVX protein expression or NOVX protein activity. The invention also includes compounds identified in the screening assays described herein.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of the membrane-bound form of a NOVX protein or polypeptide or biologically-active portion thereof. The test compounds of the invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds. *See, e.g., Lam, 1997. Anticancer Drug Design* 12: 145.

A "small molecule" as used herein, is meant to refer to a composition that has a molecular weight of less than about 5 kD and most preferably less than about 4 kD. Small molecules can be, *e.g.*, nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic or inorganic molecules. Libraries of chemical and/or biological mixtures, such as fungal, bacterial, or algal extracts, are known in the art and can be screened with any of the assays of the invention.

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt, *et al.*, 1993. *Proc. Natl. Acad. Sci. U.S.A.* 90: 6909; Erb, *et al.*, 1994. *Proc. Natl. Acad. Sci. U.S.A.* 91: 11422; Zuckermann, *et al.*, 1994. *J. Med. Chem.* 37: 2678; Cho, *et al.*, 1993. *Science* 261: 1303; Carrell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33: 2059; Carell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33: 2061; and Gallop, *et al.*, 1994. *J. Med. Chem.* 37: 1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten, 1992. *Biotechniques* 13: 412-421), or on beads (Lam, 1991. *Nature* 354: 82-84), on chips (Fodor, 1993. *Nature* 364: 555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner, U.S. Patent 5,233,409), plasmids (Cull, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 1865-1869) or on phage (Scott and Smith, 1990. *Science* 249: 386-390; Devlin, 1990. *Science* 249: 404-406; Cwirla, *et al.*, 1990. *Proc. Natl. Acad. Sci. U.S.A.* 87: 6378-6382; Felici, 1991. *J. Mol. Biol.* 222: 301-310; Ladner, U.S. Patent No. 5,233,409.).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of NOVX protein, or a biologically-active portion thereof, on the

cell surface is contacted with a test compound and the ability of the test compound to bind to a NOVX protein determined. The cell, for example, can be of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the NOVX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the NOVX protein or biologically-active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically-labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of NOVX protein, or a biologically-active portion thereof, on the cell surface with a known compound which binds NOVX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NOVX protein, wherein determining the ability of the test compound to interact with a NOVX protein comprises determining the ability of the test compound to preferentially bind to NOVX protein or a biologically-active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of NOVX protein, or a biologically-active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the NOVX protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of NOVX or a biologically-active portion thereof can be accomplished, for example, by determining the ability of the NOVX protein to bind to or interact with a NOVX target molecule. As used herein, a "target molecule" is a molecule with which a NOVX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a NOVX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A NOVX target molecule can be a non-NOVX molecule or a NOVX protein or polypeptide of the invention. In one embodiment, a NOVX target molecule is a component of a signal transduction pathway that

facilitates transduction of an extracellular signal (*e.g.* a signal generated by binding of a compound to a membrane-bound NOVX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with NOVX.

- 5 Determining the ability of the NOVX protein to bind to or interact with a NOVX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the NOVX protein to bind to or interact with a NOVX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target
- 10 molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+} , diacylglycerol, IP_3 , *etc.*), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (comprising a NOVX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell
- 15 survival, cellular differentiation, or cell proliferation.

- In yet another embodiment, an assay of the invention is a cell-free assay comprising contacting a NOVX protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to bind to the NOVX protein or biologically-active portion thereof. Binding of the test compound to the NOVX protein can
- 20 be determined either directly or indirectly as described above. In one such embodiment, the assay comprises contacting the NOVX protein or biologically-active portion thereof with a known compound which binds NOVX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NOVX protein, wherein determining the ability of the test compound to interact with
- 25 a NOVX protein comprises determining the ability of the test compound to preferentially bind to NOVX or biologically-active portion thereof as compared to the known compound.

- In still another embodiment, an assay is a cell-free assay comprising contacting NOVX protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.* stimulate or inhibit) the activity of the
- 30 NOVX protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of NOVX can be accomplished, for example, by determining the ability of the NOVX protein to bind to a NOVX target molecule by one of

the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of NOVX protein can be accomplished by determining the ability of the NOVX protein further modulate a NOVX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an
5 appropriate substrate can be determined as described, *supra*.

In yet another embodiment, the cell-free assay comprises contacting the NOVX protein or biologically-active portion thereof with a known compound which binds NOVX protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NOVX protein, wherein
10 determining the ability of the test compound to interact with a NOVX protein comprises determining the ability of the NOVX protein to preferentially bind to or modulate the activity of a NOVX target molecule.

The cell-free assays of the invention are amenable to use of both the soluble form or the membrane-bound form of NOVX protein. In the case of cell-free assays comprising the
15 membrane-bound form of NOVX protein, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of NOVX protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®,
20 Isotridecypoly(ethylene glycol ether)_n, N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl) dimethylamminiol-1-propane sulfonate (CHAPS), or 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO).

In more than one embodiment of the above assay methods of the invention, it may be desirable to immobilize either NOVX protein or its target molecule to facilitate
25 separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to NOVX protein, or interaction of NOVX protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and
30 micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, GST-NOVX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione

sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or NOVX protein, and the mixture is incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described, *supra*. Alternatively, the complexes can be dissociated from the matrix, and the level of NOVX protein binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the NOVX protein or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated NOVX protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well-known within the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with NOVX protein or target molecules, but which do not interfere with binding of the NOVX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or NOVX protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the NOVX protein or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the NOVX protein or target molecule.

In another embodiment, modulators of NOVX protein expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of NOVX mRNA or protein in the cell is determined. The level of expression of NOVX mRNA or protein in the presence of the candidate compound is compared to the level of expression of NOVX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of NOVX mRNA or protein expression based upon this comparison. For example, when expression of NOVX mRNA or protein is greater (*i.e.*, statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of

NOVX mRNA or protein expression. Alternatively, when expression of NOVX mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of NOVX mRNA or protein expression. The level of NOVX mRNA or protein expression in the cells can be
5 determined by methods described herein for detecting NOVX mRNA or protein.

In yet another aspect of the invention, the NOVX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (*see, e.g.*, U.S. Patent No. 5,283,317; Zervos, *et al.*, 1993. *Cell* 72: 223-232; Madura, *et al.*, 1993. *J. Biol. Chem.* 268: 12046-12054; Bartel, *et al.*, 1993. *Biotechniques* 14: 920-924; Iwabuchi, *et al.*, 1993. *Oncogene* 8: 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or
10 interact with NOVX ("NOVX-binding proteins" or "NOVX-bp") and modulate NOVX activity. Such NOVX-binding proteins are also involved in the propagation of signals by the NOVX proteins as, for example, upstream or downstream elements of the NOVX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for NOVX is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an
20 unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a NOVX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) that is operably linked to a transcriptional
25 regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with NOVX.

The invention further pertains to novel agents identified by the aforementioned screening assays and uses thereof for treatments as described herein.

Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. By way of example, and not of limitation, these sequences can be used to: (i) map
5 their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. Some of these applications are described in the subsections, below.

Chromosome Mapping

10 Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the NOVX sequences of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, or fragments or derivatives thereof, can be used to map the location of the NOVX genes, respectively, on a
15 chromosome. The mapping of the NOVX sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, NOVX genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the NOVX sequences. Computer analysis of the NOVX sequences can be used to rapidly select primers that do not span more than one
20 exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the NOVX sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals
25 (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but in which human cells can, the one human chromosome that contains the gene encoding the needed enzyme will be retained. By using various media, panels of hybrid cell lines can
30 be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy

mapping of individual genes to specific human chromosomes. *See, e.g., D'Eustachio, et al., 1983. Science 220: 919-924.* Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

- 5 PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the NOVX sequences to design oligonucleotide primers, sub-localization can be achieved with panels of fragments from specific chromosomes.
- 10 Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical like colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and
- 15 dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases, will suffice to get good results at a
- 20 reasonable amount of time. For a review of this technique, *see, Verma, et al., HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES* (Pergamon Press, New York 1988).

- Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to
- 25 noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

- Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such
- 30 data are found, *e.g., in McKusick, MENDELIAN INHERITANCE IN MAN*, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through

linkage analysis (co-inheritance of physically adjacent genes), described in, *e.g.*, Egeland, *et al.*, 1987. *Nature*, 325: 783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the NOVX gene, can be determined. If a mutation
5 is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete
10 sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

Tissue Typing

The NOVX sequences of the invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested with
15 one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Patent No. 5,272,057).

Furthermore, the sequences of the invention can be used to provide an alternative
20 technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the NOVX sequences described herein can be used to prepare two PCR primers from the 5'- and 3'-termini of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner,
25 can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the invention can be used to obtain such identification sequences from individuals and from tissue. The NOVX sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater
30 degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic

variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes.

- 5 Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If coding sequences, such as those of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, are used, a
10 more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

- The invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for
15 prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the invention relates to diagnostic assays for determining NOVX protein and/or nucleic acid expression as well as NOVX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder,
20 associated with aberrant NOVX expression or activity. The disorders include metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with
25 chronic diseases and various cancers. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with NOVX protein, nucleic acid expression or activity. For example, mutations in a NOVX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the
30 onset of a disorder characterized by or associated with NOVX protein, nucleic acid expression, or biological activity.

Another aspect of the invention provides methods for determining NOVX protein, nucleic acid expression or activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of NOVX in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

An exemplary method for detecting the presence or absence of NOVX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting NOVX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes NOVX protein such that the presence of NOVX is detected in the biological sample. An agent for detecting NOVX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to NOVX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length NOVX nucleic acid, such as the nucleic acid of SEQ ID NO:2*n*-1, wherein *n* is an integer between 1 and 127, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to NOVX mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting NOVX protein is an antibody capable of binding to NOVX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently-labeled secondary

antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently-labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect

5 NOVX mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of NOVX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of NOVX protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, and immunofluorescence. *In vitro* techniques for detection of

10 NOVX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of NOVX protein include introducing into a subject a labeled anti-NOVX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test

15 subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent

20 capable of detecting NOVX protein, mRNA, or genomic DNA, such that the presence of NOVX protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of NOVX protein, mRNA or genomic DNA in the control sample with the presence of NOVX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of NOVX in a

25 biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting NOVX protein or mRNA in a biological sample; means for determining the amount of NOVX in the sample; and means for comparing the amount of NOVX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect NOVX protein or nucleic

30 acid.

Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant NOVX expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with NOVX protein, nucleic acid expression or activity. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the invention provides a method for identifying a disease or disorder associated with aberrant NOVX expression or activity in which a test sample is obtained from a subject and NOVX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of NOVX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant NOVX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant NOVX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder. Thus, the invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant NOVX expression or activity in which a test sample is obtained and NOVX protein or nucleic acid is detected (*e.g.*, wherein the presence of NOVX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant NOVX expression or activity).

The methods of the invention can also be used to detect genetic lesions in a NOVX gene, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized by aberrant cell proliferation and/or differentiation. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a NOVX-protein, or the misexpression of the NOVX gene. For example,

such genetic lesions can be detected by ascertaining the existence of at least one of: (i) a deletion of one or more nucleotides from a NOVX gene; (ii) an addition of one or more nucleotides to a NOVX gene; (iii) a substitution of one or more nucleotides of a NOVX gene, (iv) a chromosomal rearrangement of a NOVX gene; (v) an alteration in the level of a messenger RNA transcript of a NOVX gene, (vi) aberrant modification of a NOVX gene, such as of the methylation pattern of the genomic DNA, (vii) the presence of a non-wild-type splicing pattern of a messenger RNA transcript of a NOVX gene, (viii) a non-wild-type level of a NOVX protein, (ix) allelic loss of a NOVX gene, and (x) inappropriate post-translational modification of a NOVX protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a NOVX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (*see, e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (*see, e.g.*, Landegran, *et al.*, 1988. *Science* 241: 1077-1080; and Nakazawa, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 360-364), the latter of which can be particularly useful for detecting point mutations in the NOVX-gene (*see*, Abravaya, *et al.*, 1995. *Nucl. Acids Res.* 23: 675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to a NOVX gene under conditions such that hybridization and amplification of the NOVX gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (*see*, Guatelli, *et al.*, 1990. *Proc. Natl. Acad. Sci. USA* 87: 1874-1878), transcriptional amplification system (*see*, Kwoh, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 1173-1177);

Q β Replicase (*see*, Lizardi, *et al.*, 1988. *BioTechnology* 6: 1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a NOVX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (*see, e.g.*, U.S. Patent No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in NOVX can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high-density arrays containing hundreds or thousands of oligonucleotide probes. *See, e.g.*, Cronin, *et al.*, 1996, *Human Mutation* 7: 244-255; Kozal, *et al.*, 1996, *Nat. Med.* 2: 753-759. For example, genetic mutations in NOVX can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, *et al.*, *supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the NOVX gene and detect mutations by comparing the sequence of the sample NOVX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert, 1977. *Proc. Natl. Acad. Sci. USA* 74: 560 or Sanger, 1977. *Proc. Natl. Acad. Sci. USA* 74: 5463. It is also contemplated that any of a variety of automated sequencing

procedures can be utilized when performing the diagnostic assays (*see, e.g., Naeve, et al., 1995. Biotechniques 19: 448*), including sequencing by mass spectrometry (*see, e.g., PCT International Publication No. WO 94/16101; Cohen, et al., 1996, Adv. Chromatography 36: 127-162; and Griffin, et al., 1993. Appl. Biochem. Biotechnol. 38: 147-159*).

5 Other methods for detecting mutations in the NOVX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes. *See, e.g., Myers, et al., 1985. Science 230: 1242*. In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type NOVX sequence with
10 potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S₁ nuclease to enzymatically digesting the mismatched regions. In other embodiments,
15 either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. *See, e.g., Cotton, et al., 1988. Proc. Natl. Acad. Sci. USA 85: 4397; Saleeba, et al., 1992. Methods Enzymol. 217: 286-295*. In
20 an embodiment, the control DNA or RNA can be labeled for detection.

 In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in NOVX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli*
25 cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches. *See, e.g., Hsu, et al., 1994. Carcinogenesis 15: 1657-1662*. According to an exemplary embodiment, a probe based on a NOVX sequence, *e.g., a* wild-type NOVX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage
30 products, if any, can be detected from electrophoresis protocols or the like. *See, e.g., U.S. Patent No. 5,459,039*.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in NOVX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids. *See, e.g., Orita, et al., 1989. Proc. Natl. Acad. Sci. USA: 86: 2766; Cotton,*
5 *1993. Mutat. Res. 285: 125-144; Hayashi, 1992. Genet. Anal. Tech. Appl. 9: 73-79.*

Single-stranded DNA fragments of sample and control NOVX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected
10 with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. *See, e.g., Keen, et al., 1991. Trends Genet. 7: 5.*

15 In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). *See, e.g., Myers, et al., 1985. Nature 313: 495.* When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of
20 high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. *See, e.g., Rosenbaum and Reissner, 1987. Biophys. Chem. 265: 12753.*

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective
25 primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. *See, e.g., Saiki, et al., 1986. Nature 324: 163; Saiki, et al., 1989. Proc. Natl. Acad. Sci. USA 86: 6230.* Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different
30 mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization; *see, e.g., Gibbs, et al.,* 5 *1989. Nucl. Acids Res.* 17: 2437-2448) or at the extreme 3'-terminus of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (*see, e.g., Prossner, 1993. Tibtech.* 11: 238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. *See, e.g.,* Gasparini, *et al., 1992. Mol. Cell Probes* 6: 1. It is anticipated that in certain embodiments 10 amplification may also be performed using *Taq* ligase for amplification. *See, e.g., Barany, 1991. Proc. Natl. Acad. Sci. USA* 88: 189. In such cases, ligation will occur only if there is a perfect match at the 3'-terminus of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

15 The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.,* in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a NOVX gene.

20 Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which NOVX is expressed may be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

25 Agents, or modulators that have a stimulatory or inhibitory effect on NOVX activity (*e.g.,* NOVX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders. The disorders include but are not limited to, *e.g.,* those diseases, disorders and conditions listed above, and more particularly include those diseases, disorders, or conditions associated with 30 homologs of a NOVX protein, such as those summarized in Table A.

In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation
5 between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of NOVX protein, expression of NOVX
10 nucleic acid, or mutation content of NOVX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons.
15 See *e.g.*, Eichelbaum, 1996, *Clin. Exp. Pharmacol. Physiol.*, 23: 983-985; Linder, 1997. *Clin. Chem.*, 43: 254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic
20 conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

25 As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome pregnancy zone protein precursor enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or
30 show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different

among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of NOVX protein, expression of NOVX nucleic acid, or mutation content of NOVX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a NOVX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of NOVX (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase NOVX gene expression, protein levels, or upregulate NOVX activity, can be monitored in clinical trials of subjects exhibiting decreased NOVX gene expression, protein levels, or downregulated NOVX activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease NOVX gene expression, protein levels, or downregulate NOVX activity, can be monitored in clinical trials of subjects exhibiting increased NOVX gene expression, protein levels, or upregulated NOVX activity. In such clinical trials, the expression or activity of NOVX and, preferably, other genes that have been implicated in, for example, a cellular proliferation or immune disorder can be used as a "read out" or markers of the immune responsiveness of a particular cell.

By way of example, and not of limitation, genes, including NOVX, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) that modulates NOVX activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of NOVX and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of NOVX or other genes. In this manner, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a NOVX protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the NOVX protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the NOVX protein, mRNA, or genomic DNA in the pre-administration sample with the NOVX protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of NOVX to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of NOVX to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant NOVX expression or activity. The disorders include but are not limited to, *e.g.*, those
 5 diseases, disorders and conditions listed above, and more particularly include those diseases, disorders, or conditions associated with homologs of a NOVX protein, such as those summarized in Table A.

These methods of treatment will be discussed more fully, below.

Diseases and Disorders

10 Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to: (i) an aforementioned peptide, or analogs,
 15 derivatives, fragments or homologs thereof; (ii) antibodies to an aforementioned peptide; (iii) nucleic acids encoding an aforementioned peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to an aforementioned peptide) that are utilized to "knockout" endogenous function of an aforementioned peptide by homologous
 20 recombination (*see, e.g.*, Capecchi, 1989. *Science* 244: 1288-1292); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between an aforementioned peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not
 25 suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, an aforementioned peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

30 Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro*

for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of an aforementioned peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, *etc.*) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, and the like).

Prophylactic Methods

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant NOVX expression or activity, by administering to the subject an agent that modulates NOVX expression or at least one NOVX activity. Subjects at risk for a disease that is caused or contributed to by aberrant NOVX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the NOVX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending upon the type of NOVX aberrancy, for example, a NOVX agonist or NOVX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein. The prophylactic methods of the invention are further discussed in the following subsections.

Therapeutic Methods

Another aspect of the invention pertains to methods of modulating NOVX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of NOVX protein activity associated with the cell. An agent that modulates NOVX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a NOVX protein, a peptide, a NOVX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more NOVX protein activity. Examples of such stimulatory agents include active NOVX protein and a nucleic acid molecule encoding NOVX that has been introduced into the cell. In another embodiment, the agent inhibits one or more NOVX protein activity. Examples of such inhibitory agents include antisense NOVX nucleic acid molecules and anti-NOVX

- antibodies. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a NOVX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., up-regulates or down-regulates) NOVX expression or activity. In another embodiment, the method involves administering a NOVX protein or nucleic acid molecule as therapy to compensate for reduced or aberrant NOVX expression or activity.
- 10 Stimulation of NOVX activity is desirable *in situations* in which NOVX is abnormally downregulated and/or in which increased NOVX activity is likely to have a beneficial effect. One example of such a situation is where a subject has a disorder characterized by aberrant cell proliferation and/or differentiation (e.g., cancer or immune associated disorders). Another example of such a situation is where the subject has a
- 15 gestational disease (e.g., preclampsia).

Determination of the Biological Effect of the Therapeutic

- In various embodiments of the invention, suitable *in vitro* or *in vivo* assays are performed to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.
- 20 In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly,
- 25 for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

Prophylactic and Therapeutic Uses of the Compositions of the Invention

- The NOVX nucleic acids and proteins of the invention are useful in potential prophylactic and therapeutic applications implicated in a variety of disorders. The disorders
- 30 include but are not limited to, e.g., those diseases, disorders and conditions listed above, and

more particularly include those diseases, disorders, or conditions associated with homologs of a NOVX protein, such as those summarized in Table A.

As an example, a cDNA encoding the NOVX protein of the invention may be useful in gene therapy, and the protein may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the invention will have efficacy for treatment of patients suffering from diseases, disorders, conditions and the like, including but not limited to those listed herein.

Both the novel nucleic acid encoding the NOVX protein, and the NOVX protein of the invention, or fragments thereof, may also be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. A further use could be as an anti-bacterial molecule (*i.e.*, some peptides have been found to possess anti-bacterial properties). These materials are further useful in the generation of antibodies, which immunospecifically-bind to the novel substances of the invention for use in therapeutic or diagnostic methods.

The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLES

Example A: Polynucleotide and Polypeptide Sequences, and Homology Data

Example 1.

The NOV1 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 1A.

Table 1A. NOV1 Sequence Analysis			
	SEQ ID NO: 1	6988 bp	
NOV1a, CG108 440-01 DNA Sequence	GAAGAGCAAGAGGCAGGCTCAGCAAATGGTTCAGCCCCAGTCCCCGGTGGCTGTCAGTCAA AGCAAGCCCGGTTGTTATGACAATGGAAAACACTATCAGATAAATCAACAGTGGGAGCGGA CCTACCTAGGTAATGTGTTGGTTTGTACTTGTTATGGAGGAAGCCGAGGTTTAACTGCGA AAGTAAACCTGAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTG GGTGACACTTATGAGCGTCCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTG GGCGAGGGAGAATAAGCTGTACCATCGCAAACCGCTGCCATGAAGGGGGTCAGTCTTACAA GATTGGTGACACCTGGAGGAGACCACATGAGACTGGTGGTTACATGTTAGAGTGTGTGTGT CTTGGTAATGGAAAAGGAGAATGGACCTGCAAGCCCATAGCTGAGAAGTGTTTTGATCATG CTGCTGGGACTTCCTATGTGGTCGGAGAAAACGTGGGAGAAGCCCTACCAAGGCTGGATGAT GGTAGATTGTACTTGCCTGGGAGAAGGCAGCGGACGCATCACTTGCACTTCTAGAAATAGA TGCAACGATCAGGACACAAGGACATCCTATAGAATTGGAGACACCTGGAGCAAGAAGGATA ATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGGCCGAGGAGAGTGGAAGTGTGA		

GAGGCACACCTCTGTGCAGACCACATCGAGCGGATCTGGCCCCCTTACCAGATGTTCTGTGCA
GCTGTTTACCAACCGCAGCCTCACCCCCAGCCTCCTCCCTATGGCCACTGTGTACAGACA
GTGGTGTGGTCTACTCTGTGGGGATGCAGTGGTTGAAGACACAAGGAAATAAGCAAATGCT
TTGCACGTGCCTGGGCAACGGAGTCAGCTGCCAAGAGACAGCTGTAAACCCAGACTTACGGT
GGCAACTTAAATGGAGAGCCATGTGTCTTACCATTACCTACAATGGCAGGACGTTCTACT
CCTGCACCACGGAAGGGCGACAGGACGGACATCTTGGTGCAGCACAACTTCGAATTATGA
GCAGGACCAGAAATACTCTTTCTGCACAGACCACACTGTTTGGTTTCAGACTCAAGGAGGA
AATTCCAATGGTGCCTTGTGCCACTTCCCCTTCTATACAACAACCACAATTACACTGATT
GCACTTCTGAGGGCAGAAGAGACAACATGAAGTGGTGTGGGACCACACAGAACTATGATGC
CGACCAGAAGTTTGGGTCTGCCCCATGGCTGCCACGAGGAAATCTGCACAACCAATGAA
GGGTTCATGTACCGCATTGGAGATCAGTGGGATAAGCAGCATGACATGGGTACATGATGA
GGTGCACGTGTGTTGGGAATGGTCGTGGGAATGGACATGCATTGCCTACTCGCAACTTCG
AGATCAGTGCATTGTTGATGACATCACTTACAATGTGAACGACACATTCCACAAGCGTCAT
GAAGAGGGGCACATGCTGAACGTACATGCTTCGGTCAGGGTCGGGGCAGGTGGAAGTGTG
ATCCCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTATCAAATTGGAGATTTCATG
GGAGAAGTATGTGCATGGTGTGATACAGTACAGTGTCTACTGCTATGGCCGTGGCATTGGGGAG
TGGCATTGGCCAACCTTTACAGACCTATCCAAGCTCAAGTGGTCCTGTGGAAGTATTTATCA
CTGAGACTCCGAGTCAGCCCACTCCACCCCATCCAGTGAATGCACCACAGCCATCTCA
CATTTCCAAGTACATTCTCAGGTGGAGACCTAAAAATTCTGTAGGCCGTTGGAAGGAAGCT
ACCATACCAGGCCACTTAAACTCTACACCATCAAAGGCTGAAGCCTGGTGTGGTATACG
AGGGCCAGCTCATCAGCATCCAGCAGTACGGCCACCAAGAAGTGACTCGCTTTGACTTCAC
CACCACCAGCACCAGCACACTGTGACCAGCAACACCGTGACAGGAGAGACGACTCCCTTT
TCTCCTCTTGTGGCCACTTCTGAATCTGTGACCAGAAATCACAGCCAGTAGCTTTGTGGTCT
CCTGGGTCTCAGCTTCCGACACCGTGTGCGGATTCCGGGTGGAATATGAGCTGAGTGAGGA
GGGAGATGAGCCACAGTACCTGGATCTTCCAAGCACAGCCACTTCTGTGAACATCCCTGAC
CTGCTTCTTGCCGAAAATACATTGTAATGTCTATCAGATATCTGAGGATGGGGAGCAGA
GTTGATCCTGTCTACTTCACAAACAACAGCGCCTGATGCCCTCCTGACCCGACTGTGGA
CCAAGTTGATGACACCTCAATTGTGTTCGCTGGAGCAGACCCAGGCTCCCATCACAGGG
TACAGAATAGTCTATTGCGCATCAGTAGAAGGTAGCAGCACAGAACTCAACCTTCTGAAA
CTGCAAACTCCGTCAACCTCAGTGACTTGCAACCTGGTGTTCAGTATAACATCACTATCTA
TGCTGTGGAAGAAAATCAAGAAAGTACACCTGTGTGTCATTCAACAAGAAACCACTGGCACC
CCACGCTCAGATACAGTGCCCTCTCCCAGGGACCTGCAGTTTGTGGAAGTGACAGACGTGA
AGGTCAACCATCATGTGGACACCGCTGAGAGTGCAGTGACCGGTACCGTGTGGATGTGAT
CCCCGTCAACCTGCCTGGCGAGCAGGGCAGAGGCTGCCCATCAGCAGGAACACCTTTGCA
GAAGTCACCGGGCTGTCCCTGGGGTCACCTATTACTTCAAAGTCTTTGCAGTGAGCCATG
GGAGGGAGAGCAAGCCTCTGACTGCTCAACAGACAACCAAACTGGATGCTCCCACTAACCT
CCAGTTTGTCAATGAAACTGATTCTACTGTCTGGTGAGATGGACTCCACCTCGGGCCAG
ATAACAGGATACCGACTGACCGTGGGCCTTACCCGAAGAGGCCAGCCAGGCAGTACAATG
TGGGTCCCTCTGTCTCCAAGTACCCCTGAGGAATCTGCAGCCTGCATCTGAGTACACCGT
ATCCCTCGTGGCCATAAAGGGCAACCAAGAGAGCCCCAAAGCCACTGGAGTCTTTACCACA
CTGCAGCCTGGGAGCTCTATTCCACCTTACAACACCGAGGTGACTGAGACCACCATCGTGA
TCACATGGACGCTGTCTCAAGAATTGGTTTAAAGCTGGGTGTACGACCAAGCCAGGGAGG
AGAGGCACCACGAGAAGTGACTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCA
GGAGTAGAATACGTCTACACCATCCAAGTCTGAGAGATGGACAGGAAAGAGATGCGCCAA
TTGTAAACAAAGTGGTGACACCATGTCTCCACCAACAACTTGCACTCTGGAGGCAAAACC
TGACACTGGAGTGCTCACAGTCTCCTGGGAGAGGAGCACCACCCAGACATTACTGGTTAT
AGAATTACCACAACCCCTACAAACGGCCAGCAGGGAAATCTTTGGAAGAAGTGGTCCATG
CTGATCAGAGCTCCTGCACCTTTTGATAACCTGAGTCCCGGCTGGAGTACAATGTCAGTGT
TTACACTGTCAAGGATGACAAGGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTT
CCTCCTCCCACTGACCTGCGATTACCAACATTGGTCCAGACACCATGCGTGTACCTGGG
CTCCACCCCATCCATTGATTTAACCAACTTCTGGTGCCTTACTCACCTGTGAAAAATGA
GGAAGATGTTGCAGAGTTGTCAATTTCTCCTTCAGACAATGCAGTGGTCTTAACAAATCTC
CTGCCCTGTACAGAATATGTAGTGAGTGTCTCCAGTGTCTACGAACAACATGAGAGCACAC
CTCTTAGAGGAAGACAGAAAACAGGTCTTGATTCCCCAACTGGCATTGACTTTTCTGATAT
TACTGCCAACTCTTTTACTGTGCACTGGATTGCTCCTCGAGCCACCATCACTGGCTACAGG
ATCCGCCATCATCCCGAGCACTTCAGTGGGAGACCTCGAGAAGATCGGGTGCCCACTCTC
GGAATTCCATCACCCCTACCAACCTCACTCCAGGCACAGAGTATGTGGTCAGCATCGTTGC
TCTTAATGGCAGAGAGGAAAGTCCCTTATTGATTGGCCAACAATCAACAGTTTCTGATGTT
CCGAGGGACCTGGAAGTTGTTGCTGCGACCCCCACCAGCCTACTGATCAGCTGGGATGCTC
CTGCTGTACAGTGAGATATTACAGGATCACTTACGGAGAAACAGGAGGAAATAGCCCTGT

	CCAGGAGTTCAGTGTGCTGGGAGCAAGTCTACAGCTACCATCAGCGGCCTTAAACCTGGA GTTGATTATACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCCGCAAGCAGCA AGCCAATTTCCATTAAATTACCGAACAGAAATTGACAAACCATCCCAGATGCAAGTGACCGA TGTTTCAGGACAACAGCATTAGTGTCAAGTGGCTGCCTTCAAGTTCCTGTTACTGGTTAC AGAGTAACCACCACTCCCAAAAATGGACCAGGACCAACAAAACTAAAATGTCAGGTCAG ATCAAAACAGAAATGACTATTGAAGGCTTGCAGCCACAGTGGAGTATGTGGTTAGTGTCTA TGCTCAGAATCCAAGCGGAGAGAGTCAGCCTCTGGTTCAGACTGCAGTAACCAACATTGAT CGCCCTAAAGGACTGGCATTCACTGATGTGGATGTCGATTCCATCAAATTGCTTGGGAAA GCCACAGGGGCAAGTTCCAGGTACAGGGTGACCTACTCGAGCCCTGAGGATGGAATCCA TGAGCTATTCCCTGCACCTGATGGTGAAGAAGACACTGCAGAGCTGCAAGGCCCTCAGACCG GGTCTGAGTACACAGTCAGTGTGGTTGCCTTGCACGATGATATGGAGAGCCAGCCCTGA TTGGAACCCAGTCCACAGCTATTCTGCACCAACTGACCTGAAGTTCAGTCAAGTCAACACC CACAAGCCTGAGCGCCAGTGGACACCACCAATGTTCACTCACTGGATATCGAGTGCAGG GTGACCCCCAAGGAGAAGACCCGACCAATGAAAGAAATCAACCTTCTCCTGACAGCTCAT CCGTGGTTGTATCAGGACTTATGGTGGCCACCAATATGAAGTGAAGTGTCTATGCTCTTAA GGACACTTTGACAAGCAGACAGCTCAGGGTGTGTCACTCACTCTGGAGAATGTCAGCCCCA CCAAGAAGGGCTCGTGTGACAGATGCTACTGAGACCACCATCACCATTAGCTGGAGAACCA AGACTGAGACGATCACTGGCTTCCAAGTTGATGCCGTTCCAGCCAATGGCCAGACTCCAAT CCAGAGAACCATCAAGCCAGATGTGAGAAGTACACCATCACAGGTTTACAACCAGGCACT GACTACAAGATCTACCTGTACACCTTGAATGACAATGCTCGGAGCTCCCCTGTGGTCTATCG ACGCTTCCACTGCCATTGATGCACCATCCAACCTGCGTTTCTGGCCACACACCCAATTTC CTTGCTGGTATCATGGCAGCCGCCAGTGCAGGATTACCGGCTACATCATCAAGTATGAG AAGCCTGGGTCTCTCCAGAGAAGTGGTCCCTCGGCCCGCCCTGGTGTACAGAGGGCTA CTATTACTGGCCTGGAACCGGGAACCGAATATACAATTTATGTCATTGCCCTGAAGAATAA TCAGAAGAGCGAGCCCTGATTGGAAGGAAAAAGACAGACGAGCTTCCCAACTGGTAACC CTTCCACACCCCAATCTTATGGACAGAGATCTTGGATGTTCTTCCACAGTTCAAAGA CCCCTTTCGTCACCCACCTGGGTATGACACTGGAATGGTATTACAGCTTCTGGCACTTC TGGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGAACATGGTTTTAGGCGGACC ACACCGCCCAACAACGGCCACCCCATAGGCATAGGCCAAGACCATACCCGCCGAATGTAG GACAAGAAGCTCTCTCTCAGACAACCATCTCATGGGCCCATTCAGGACACTTCTGAGTA CATCATTTTCATGTCATCTGTGGCACTGATGAAGAACCCTTACAGTTCAAGGTTCTTGGA ACTTCTACAGTGCCACTCTGACAGGCCTCACCAGAGGTGCCACCTACAACATCATAGTGG AGGCACTGAAAGACCAGCAGAGGCATAAGGTTCCGGGAAGAGGTTGTTACCGTGGGCAACTC TGTCACGAAGGCTTGAACCAACCTACGGATGACTCGTGCTTTGACCCCTACACAGTTTCC CATTATGCCGTTGGAGATGAGTGGGAACGAATGTCTGAATCAGGCTTTAACTGTTGTGCC AGTGCTTAGGCTTTGGAAGTGGTCATTTAGATGTGATTATCTAGATGGTGCCATGACAA TGGTGTGAACATAAGATTGGAGAGAAGTGGGACCGTCAGGGAGAAAATGGCCAGATGATG AGCTGCACATGTCTTGGGAACGGAAAAGGAGAATTCAAGTGTGACCCCTCATGAGGCAACGT GTTACGATGATGGGAAGACATACCACGTAGGAGAACAGTGGCAGAAGGAATATCTCGGTGC CATTTGCTCTGCACATGCTTTGGAGGCCAGCGGGGCTGGCGCTGTGACAACGTCCGCAG CCTGGGGGTGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTTACAACAGTATTCTCAGA GATACCATCAGAGAACAAACATAATGTTAATTGCCCAATTGAGTGCTTCATGCCTTTAGA TGACAGGCTGACAGAGAAGATTCCCGAGAGTAA		
	ORF Start: ATG at 26		ORF Stop: TAA at 6986
	SEQ ID NO: 2	2320 aa	MW at 255732.8kD
NOVIa,CG108 440-01 Protein Sequence	MVQPQSPVAVSQSKPGCYDNGKHYQINQQWERTYLGVLVCTCYGGSRGFNCSKPEAET CFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRISCTIANRCHEGGQSYKIGDTWRRP HETGGYMLECVCLGNGKGEWTKPIAEKCFDHAAGTSYVVGETWEKPYQGWMMVDCTCLGE GSGRITCTSRNRCNDQDTRTSYRIGDTSKKNRGNLLQCICITGNRGGEWKERHTSVQTT SSGSGPFTDVRAAVYQPPHPQPPPYGHCVTDGSGVVYSVGMQWLKTQGNKQMLCTCLNGV SCQETAVTQTYGGNLNGEPCVLPFTYNGRTFYSCCTTEGRQDGLWCSTTSNYEQDQKYSFC TDHTVLVQTQGGNSNGALCHFPFLYNNHNYTDCTSEGRDNDMKWCGTQNYDADQKFGFCP MAAHEEICTTNEGVMYRIGDQWDKQHDMMMRCTCVGNRGGEWTCIAYSQLRDQCIIVDDI TYNVNDTFHKRHEEGHMLNCTCFGQGRGRWKCDPVDQCDSETGTFYQIGDSWEKYVHGV YQCYCYGRGIGEWHCQPLQTYPSSSGPVVFITETPSQPNSHPIQWNAPQPSHISKYILRW RPKNSVGRWKEATIPGHLNSYTIKGLKPGVVYEGQLISIQQYGHQEVTRFDFTTSTSTPV TSNTVTGETTTPFSPLVATSESVEITASSFVVSWSASDTVSGFRVEYELSEEGDEPQYLD LPSTATSVNIPDLLPGRKYIVNVYQISEDEQSLILSTSQTTPADAPPDPTVDQVDDTSIV		

	<p>VRWSRPQAPITGYRIVYSPSVEGSSTELNLPETANSVTLSDLQPGVQYNNITIYAVEENQES TPVVIQQUETTGTTPRSPTVPSRDLQFVEVTDVKVTIMWTPPESAVTGYRVDVIPVNLPGEH GQRLPISRNTFAEVTGLSPGVYTYFKVFAVSHGRESKPLTAQQTTKLDAPTNLQFVNETDS TVLVRWTPPRAQITGYRLTVGLTRRGQPRQYNVGPVSQYPLRNLPASEYTVSLVAIKGN QESPKATGVFTTLQPGSSIPPYNTEVTETTIVITWTAPRIGFKLGVRPSQGGAEAPREVS DSGSIVVSGLTGPGVEYVYTIQVLRDQGERDAPIVNKVVTPLSPTNLHLEANPDGTGLTVS WERSTTPDITGYRITTTPTNGQQNSLEEVVHADQSSCTFDNLSPGLEYNVSVYTVKDDKE SVPISDTIIPAVPPPTDLRFTNIGPDMRVTWAPPPSIDLTNLFVRYSVPVKNEDVAELSI SPSDNAVVLTNLLPGTEYVVSVSVEYQHESTPLRGRQKTGLDSPTGIDFSDITANSFTVH WIAPRATITGYRIRHHEHFGSRPREDRVPHSRNSITLTNLTGTEYVVSIVALNGREES LLIGQOSTVSDVPRDLEVVAAPTSLLSISWDAPAVTVRYRITYGETGGNSPVQEFVTPGS KSTATISGLKPGVDYTTIVYAVTGRGDSPASSKPISINRYTEIDKPSQMQVTDVQDNSISV KWLPSSTPVTGYRVTTTPKNGPGPTKTKTAGPDQTEMTIEGLQPTVEYVVSVAQNPSGES QPLVQTAVTNIDRPKGLAFTDQVDSIKIAWESPGQVSRVRYTYSSPEDGHELFPAAPDG EEDTAELQGLRPGSEYTVSVVALHDDMESQPLIGTQSTAIAPATDLKFTQVTPSLSAQWT PPNVQLTGYRVRVTPKEKTGPMKEINLAPDSSSVVSGLMVATKYEVSVYALKDITLTSRPA QGVVTTLENVSPRRARVDATETTITISWRTKTETITGFQVDAVPANGQTPQRTIKPDV RSYTTITGLQPGTDYKIYLYTLNDNARSSPVVIDASTAIDAPSNLRLFLATTPNSLLVSWQPP RARITGYIIKYEKPGSPPREVVRPRPGVTEATITGLEPGTEYTIIVIALKNNQKSEPLIG RKKTDELQPLVTLPHENLHGPEILDVPSTVQKTPFVTHPGYDTGNGIQLPGTSGQQPSVGQ QMIFEEHGFRTTPTTATPIRHRPRPYPPNVGQEAQSQTISWAPFQDTSEYIISCHPVG TDEEPLQFRVPGTSTSATLTGLTRGATYNIIVEALKDQQRHKVREEVTVGNSVNEGLNQ TDDSCFDPTYVSHYAVGDEWERMESGFKLLCQCLGFGSGHFRCDSSRWCHDNGVNYKIGE KWDRQGENGQMMSCCLGNGKGEFKCDPHEATCYDDGKTYHVGEQWQKEYLGAICSTCFG GQRGWRCDCNRRPGGEPSPGTTGQSYNQYSQRYHQRTNTNVNCPICFMPDLVQADREDS RE</p>
	<p>SEQ ID NO: 3 7361 bp</p>
NOV1b,CG108 440-02 DNA Sequence	<p>TCAACATGCTTAGGGGTCCGGGGCCCGGGCTGCTGCTGCTGGCCGTCCAGTGCCCTGGGGAC AGCGGTGCCCTCCACGGGAGCCTCGAAGAGCAAGAGGCAGGCTCAGCAAATGGTTCAGCCC CAGTCCCCGGTGGCTGTCAGTCAAAGCAAGCCCGTTGTTATGACAAATGGAAAACACTATC AGATAAATCAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTTGTACTTGTATGG AGGAAGCCGAGGTTTTAACTGCGAGAGTAACCTGAAGCTGAAGAGACTTGCTTTGACAAG TACACTGGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCT GGGACTGTACCTGCATCGGGGCTGGGCGAGGGAGAATAAGCTGTACCATCGCAAACCGCTG CCATGAAGGGGGTCAGTCTACAAGATTGGTGACACCTGGAGGAGACCATGAGACTGGT GGTTACATGTAGAGTGTGTGTCTTGGTAATGGAAGGAGAATGGACCTGCAAGCCCA TAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTGGTCGGAGAAACGTGGGA GAAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAAAGGCAGCGGACGC ATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTATAGAATTG GAGACACCTGGAGCAAGAAGGATAATCGAGGAAACCTGCTCCAGTGACATCGACAGGCAA CGGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCGAGCGGATCT GGCCCCCTTACCGATGTTCTGTGCAGCTGTTTACCAACCGCAGCCTCACCCCCAGCCTCCTC CCTATGGCCACTGTGTACAGACAGTGGTGTGGTCTACTCTGTGGGGATGCAGTGGCTGAA GACACAAGGAAATAAGCAAATGCTTTCACGTGCCTGGGCAACGGAGTCAGTGCCAAGAG ACAGCTGTAACCCAGACTTACGGTGGCAACTCAAATGGAGAGCCATGTGTCTTACCATTCA CCTACAATGGCAGGACGTTCTACTCTGCACCACAGAAGGGCGACAGGACGACATCTTTG GTGCAGCACAACTTCGAATTATGAGCAGGACCAGAAATACTCTTCTGCACAGACCACACT GTTTTGGTTCAGACTCGAGGAGGAAATCCAATGGTGCCTTGTGCCACTTCCCCCTCCTAT ACAACAACCACAATTACACTGATTGCACTTCTGAGGGCAGAAGAGACAACATGAAGTGGTG TGGGACCACACAGAACTATGATGCCGACAGAAAGTTTGGGTCTGCCCCATGGCTGCCAC GAGGAAATCTGCACAACCAATGAAGGGGTGATGTACCGCATTGGAGATCAGTGGGATAAGC AGCATGACATGGGTACATGATGAGGTGCACGTGTGTTGGGAATGGTCGTGGGGAATGGAC ATGCATTGCCTACTCGCAGCTTCGAGATCAGTGCATTGTTGATGACATCACTTACAATGTG AACGACACATTCCACAAGCGTCATGAAGAGGGGCACATGCTGAAGTGTACATGCTTCGGTC AGGGTCGGGGCAGGTGGAAGTGTGATCCCGTCGACCAATGCCAGGATTACAGAGCTGGGAC GTTTTATCAAATTGGAGATTGATGGGAGAAGTATGTGATGGTGTGATGATACAGATACCGTCTAC TGCTATGGCCGTGGCATTGGGGAGTGGCATTGCCAACCTTTACAGACCTATCCAAGCTCAA GTGGTCCTGTGCAAGTATTTATCACTGAGACTCCGAGTCAGCCCACTCCCACCCATCCA GTGGAATGCACCACAGCATCTCACATTTCCAAGTACATTCTCAGGTGGAGACCTAAAAAT</p>

TCTGTAGGCCGTTGGAAGGAAGCTACCATACCAGGCCACTTAAACTCCTACACCATCAAAG
GCCTGAAGCCTGGTGTGGTATACGAGGGCCAGCTCATCAGCATCCAGCAGTACGGCCACCA
AGAAGTGACTCGCTTTGACTTCACCACCACCAGCACCAGCACACCTGTGACCAGCAACACC
GTGACAGGAGAGACGACTCCCTTTTCTCCTCTTGTGGCCACTTCTGAATCTGTGACCGAAA
TCACAGCCAGTAGCTTTGTGGTCTCCTGGGTCTCAGCTTCCGACACCGTGTCCGGATTCCG
GGTGAATATGAGCTGAGTGAGGAGGGAGATGAGCCACAGTACCTGGATCTTCCAAGCACA
GCCACTTCTGTGAACATCCCTGACCTGCTTCTGGCCGAAAATACATTGTAAATGTCTATC
AGATATCTGAGGATGGGAGCAGAGTTTGATCCTGTCTACTTCACAAACAACAGCGCCTGA
TGCCCNCCCTGACCCGACTGTGGACCAAGTTGATGACACCTCAATTGTTGTTCTGCTGGAGC
AGACCCAGGCTCCCATCACAGGGTACAGAATAGTCTATTGCGCATCAGTAGAAGGTAGCA
GCACAGAATCAACCTTCTGAACTGCAAACTCCGTACCCCTCAGTGACTTGAACCTGG
TGTTGAGTATAACATCACTATCTATGCTGTGGAAGAAAATCAAGAAAGTACACCTGTTGTC
ATTCAACAAGAAACCACTGGCACCCACGCTCAGATACAGTGCCTCTCCCAGGGACCTGC
AGTTTGTGGAAGTGACAGACGTGAAGGTCACCATCATGTGGACACCGCCTGAGAGTGACGT
GACCGGCTACCGTGTGGATGTGATCCCGTCAACCTGCCTGGCGAGCAGGGCAGAGGCTG
CCCATCAGCAGGAACACCTTTGCAGAAAGTACCGGGCTGTCCCTGGGGTCACTTATTACT
TCAAAGTCTTGCAGTGAGCCATGGGAGGGAGAGCAAGCCTCTGACTGCTCAACAGACAAC
CAAACTGGATGCTCCCACTAACCTCCAGTTTGTCAATGAACTGATTCTACTGTCTGGTG
AGATGGACTCCACCTCGGGCCAGATAACAGGATACCGACTGACCGTGGGCCTTACCCGAA
GAGGNCAGCCAGGCAGTACAATGTGGGTCCCTCTGTCTCAAGTACCCNCTGAGGAATCT
GCAGCCTGCATCTGAGTACACCGTATCCCTCGTGGCCATAAAGGGCAACCAAGAGAGCCCC
AAAGCCACTGGAGTCTTTACCACACTGCAGCCTGGGAGCTCTATTCCACCTTACAACACCG
AGGTGACTGAGACCACCATTTGTGATCACATGGACGCCTGCTCCAAGAATTGGTTTTAAGCT
GGGTGTACGACCAAGCCAGGGAGGAGAGGCCACCACGAGAAGTGACTTCAGACTCAGGAAGC
ATCGTTGTGTCCGGCTTGACTCCAGGAGTAGAATACGTCTACACCATCCAAGTCTGAGAG
ATGGACAGGAAGAGATGCGCCAATTGTAAACAAAGTGGTGACACCATTTGTCTCCACCAAC
AAACTTGCATCTGGAGGCAAAACCTGACACTGGAGTGCTCACAGTCTCCTGGGAGAGGAGC
ACCACCCAGACATTACTGGTTATAGAATTACCACAACCCCTACAAACGGCCAGCAGGGAA
ATTCTTTGGAAGAAGTGGTCCATGCTGATCAGAGCTCCTGCACCTTTTGATAACCTGAGTCC
CGGCCTGGAGTACAATGTGAGTGTACACTGTCAAGGATGACAAGGAAAGTGTCCCTATC
TCTGATACCATCATCCAGAGGTGCCCAACTCACTGACCTAAGCTTTGTTGATATAACCG
ATTCAAGCATCGGCCTGAGGTGGACCCCGCTAAACTCTTCCACCATATTGGGTACCGCAT
CACAGTAGTTGCGGCAGGAGAAGGTATCCCTATTTTGAAGATTTTGTGGACTCCTCAGTA
GGATACTACACAGTCACAGGGCTGGAGCCGGGCATTGACTATGATATCAGCGTTATCACTC
TCATTAATGGCGCGAGAGTGCCCTACTACACTGACACAACAAACGGCTGTTCTCTCTCC
CACTGACCTGCGATTACCAACATTGGTCCAGACACCATGCGTGTCACTGGGCTCCACCC
CCATCCATTGATTTAACCAACTTCTGGTGCCTTACTCACTGTGAAAATGAGGAAGATG
TTGCAGAGTTGTCAATTTCTCCTTCAGACAATGCAGTGGTCTTAACAAATCTCCTGCCTGG
TACAGAATATGTAGTGAGTGTCTCCAGTGTCTACGAACAACATGAGAGCACACCTCTTAGA
GGAAGACAGAAAACAGGTCTTGATTCCCCAACTGGCATTGACTTTTCTGATATTACTGCCA
ACTCTTTTACTGTGCACTGGATTGCTCCTCGAGCCACCATCACTGGCTACAGGATCCGCCA
TCATCCCGAGCACTTCAGTGGGAGACCTCGAGAAGATCGGGTGCCCACTCTCGGAATTCC
ATCACCTCACCAACCTCACTCCAGGCACAGAGTATGTGGTCAGCATCGTTGCTCTTAATG
GCAGAGAGGAAAGTCCCTTATTGATTGGCCAACAATCAACAGTTTCTGATGTTCCGAGGGA
CCTGGAAGTTGTTGCTGCGACCCCAACAGCCTACTGATCAGCTGGGATGCTCCTGCTGTC
ACAGTGAGATATTACAGGATCACTTACGGAGAAAACAGGAGGAAATAGCCCTGTCCAGGAGT
TCACTGTGCCTGGGAGCAAGTCTACAGCTACCATCAGCGGCCTTAAACCTGGAGTTGATTA
TACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCGCAAGCAAGCAATTT
TCCATTAATTACCGAACAGAAATTGACAAACCATCCCAGATGCAAGTGACCGATGTTTCAAG
ACAACAGCATTAGTGTCAAGTGGCTGCCTTCAAGTTCCCTGTTACTGGTTACAGAGTAAC
CACCCTCCCAAAAATGGACCAGGACCAACAAAACCTAAAACTGCAGGTCCAGATCAAACA
GAAATGACTATTGAAGGCTTGACAGCCACAGTGGAGTATGTGGTTAGTGTCTATGCTCAGA
ATCCAAGCGAGAGAGTCAAGCCTTGTTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT
AGGACTGGCATTCACTGATGTGGATGTGATTCCATCAAAATTGCTTGGGAAAGCCACAG
GGGCAAGTTTCCAGGTACAGGGTGACCTACTCGAGCCCTGAGGATGGAATCCATGAGCTAT
TCCCTGCACCTGATGGTGAAGAAGACACTGCAGAGCTGCAAGGCCCTCAGACCGGGTTCTGA
GTACACAGTCAGTGTGGTTGCCTTGCACGATGATATGGAGAGCCAGCCCTGATTGGAACC
CAGTCCACAGCTATTCTGCAACCACTGACCTGAAGTTCACTCAGTGCACACCAACAGGCC
TGAGCGCCAGTGGACACCAACCAATGTTCAAGTCACTGGATATCAGAGTGGGGTGACCC
CAAGGAGAAGACCGGACCAATGAAAGAAATCAACCTTGCTCCTGACAGCTCATCCGTGGTT

	GTATCAGGACTTATGGTGGCCACCAAATATGAAGTGAGTGTCTATGCTCTTAAGGACACTT TGACAAGCAGACCAGCTCAGGNGTTGTCAACCACTCTGGAGAATGTCAGCCCACCAAGAAG GGTCGTGTGACAGATGCTACTGAGACCACCATCACCATTAGCTGGAGAACCAAGACTGAG ACGATCACTGGCTTCCAAGTTGATGCCGTTCCAGCCAATGGCCAGACTCCAATCCAGAGAA CCATCAAGCCAGATGTCAGAAGCTACACCATCACAGGTTTACAACCAGGCACTGACTACAA GATCTACCTGTACACCTGAATGACAATGCTCGGAGCTCCCCTGTGGTCATCGACGCCTCC ACTGCCATTGATGCACCATCCAACCTGCGTTTCTGGCCACCACACCCAATTCTTGTCTGG TATCATGGCAGCCGCCACGTGCCAGGATTACCGGCTACATCATCAAGTATGAGAAGCCTGG GTCTCCTCCCAGAGAAGTGGTCCCTCGGCCCCGCCCTGGTGTACAGAGGCTACTATTACT GGCTTGAACCGGGAACCGAATATACAATTTATGTCATTGCCCTGAAGAATAATCAGAAGA GCGAGCCCCTGATTGGAAGGAAAAAGACAGGATGGTGCCATGACAATGGTGTGAACACAA GATTGGAGAGAAGTGGGACCGTCAGGGAGAAAATGGCCAGATGATGAGCTGCACATGTCTT GGGAACGGAAAAGGAGAATTCAAGTGTGACCCTCATGAGGCAACGTGTTATGATGATGGGA AGACATACCAGTAGGAGAACAGTGGCAGAAGGAATATCTCGGTGCCATTTCTCCTGCAC ATGCTTTGGAGGCCAGCGGGGCTGGCGCTGTGACAACCTGCCGAGACCTGGGGGTGAACCC AGTCCCGAAGGCACTACTGGCCAGTCTACAACCAAGTATCTCAGAGATACCATCAGAGAA CAAACACTAATGTTAATTGCCAATTGAGTGTCTCATGCCTTTAGATGTACAGGCTGACAG AGAAGATTCCCGAGAGTAAATCATCTTTCCAATCCAGAGGAACAAGCATGTCTCTGCCA AGATCCATCTAACTGGAGTGATGTTAGCAGACCCAGCTTAGAGTCTTCTTCTTTCTTA AGCCCTTTGCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCACAGCTTCTCCAAGCATC ACCCTGGGAGTTTCTGAGGGTTTCTCATAAATGAGGGCTGCACATTGCCGTGTTCTGCTT CGAAGTATTCAATACCGCTCAGTATTTTAAATGAAGTGATTCTAAGATTGTTTGGGATC AATAGGAAAGCATATGCAGCCAACCAAGATGCAAATGTTTGAATGATATGACCAAAATT TTAAGTAGGAAAGTCACCCAAACACTTCTGCTTTCACCTAAGTGTCTGGCCCGCAATACTG TAGGAACAAGCATGATCTTGTTACTGTGATATTTTAAATATCCACAGTACTCACTTTTCC AAATGATCCTAGTAATTGCCTAGAAATATCTTCTCTTACCTGTTATTTATCAATTTTCC CAGTATTTTATACGGAAGAAAATGTTATTGAAAACACTTAGTATGCAGTTGATAAGAGGAA TTTGGTATAATTATGGTGGGTGATTATTTTTTATACTGTATGTGCCAAAGCTTTACTACTG TGGAAAGACAACCTGTTTAAATAAAGATTTACATTCCACAA		
	ORF Start: at 3		ORF Stop: at 6663
	SEQ ID NO: 4	2220 aa	MW at 243994.0kD
NOV1b,CG108 440-02 Protein Sequence	MLRGPGLGLLLLAVQCLGTAVPSTGASKSKRQAQMQVQPSVAVVSQSKPGCYDNGKHYQI NQWERTYLGNALVCTCYGSGRGFNCESKPEABETCFDKYTGNTYRVGDTYERPKDSMIWD CTCIGAGRGRISCTIANRCHEGGQSYKIGDTWRRPHETGGYMLECVCLNGKGWETCKPIA EKCFDHAAGTSYVVGETWEKPYQGWMMVDCTCLGEGSGRITCTSRNRCDQDRTSYRIGD TWSKKDNRGNLLQICITGNGRGEWKCEHRTSVQTTSSGSGPFTDVRAAVYQPPHPQPPY GHCVTDSGVVSVGMQWLKTQGNKQMLCTCLGNGVSCQETAVTQTYGGNSNGEPCVLPFTY NGRTFYSCTEGRQDGLWCSTTSNYEQDQKYSFCTDHTVLVQTRGNSNGALCHFPFLYN NHNYTDCTSEGRDNMKWCCTTQNYDADQKFGFCPMAAHEEICTNEGVMYRIGDQWDQKH DMGHMMRCTCVNGRGEWTCIAYSQLRDQCIVDDITYNVNDTFHHRHEEGHMLNCTCFGGQ RGRWKCDPVDQCQDSETGTFYQIGDSWEKYVHGVRVYQCYCYGRGIGEWHCQLQTYPSSSG PVEVFITETPSQPNSHPIQWNAPOPSHISKYILRWRPKNSVGRWKEATIPGHLNSYTIKGL KPGVVYEGQLISIQYGHQEVTRFDFTTSTSTPVTSTNTVTGETTFFSPLVATSESVTEIT ASSFVVSWSASDVTSGFRVEYELSEEGDEPQYLDLPSTATSVNIPDLLPGRKYIVNVYQI SEDGEQSLILSTSQTAPDAPPDPTVDQVDDTSIVVRSRPPQAPITGYRIVYSPSVEGSST ELNLPETANSVTLSDLQPGVQYNITIYAVEENQESTPVVIQETTGTPRSDTVPSPRDLQF VEVTDVKVTIMWTPPESAVTGYRVDVIPVNLPGEHGQRLPISRNTFAEVTGLSPGVTYYPK VFAVSHGRESKPLTAQQTTLKDAPTNLQFVNETDSTVLVRWTPPRAQITGYRLTVGLTRRG QPRQYNVGPVSVKYPLRNLPASEYTVSLVAIKGNQESPKATGVFTTLQPGSSIPPYNTEV TETTIVITWTPAPRIGFKLGVRPSQGEAPREVTSDSGSIVVSGLTPGVEVYVYTIQVLRDG QERDAPIVNKVVTPPLSPPTNLHLEANPDTGVLTVSWERSTTDPITGYRITTTPTNGQQNS LEEVVHADQSSCTFDNLSPGLEYNVSVYTVKDDKESVPISDTIIPVLPQLTDLFSVDITDS SIGLRWTPLNSSTIIGYRITVVAAGEGPIFEDFVDSSVGYYTVTGLEPGIDYDISVITLI NGGESAPTTLTQQTAVPPPTDLRFTNIGPDTMRVTWAPPSIDLNLFVRYSPVKNEEDVA ELISIPSDNAVVLNLLPGTEYVVSVSVEQHESTPLRGRQKTGLDSTGIDFSDITANS FTVHWIAPRATITGYRIRHHPEHFSGRPREDRVPHSRNSITLNLTPGTEYVVSIVALNGR EESPLLIQQSTVSDVPRDLEVAATPTSLISWDAPAVTVRYRYITYGETGNSPVQEFT VPGSKSTATISGLKPGVDYITITVYAVTGRGDSPASSKPISINYRTEIDKPSQMQVTDVQDN		

	SISVKWLPSSSPVTGYRVTTTPKNGPGPTKTKTAGPDQTEMTIEGLQPTVEYVVSVAQNP SGESQPLVQTAVTNIDRPKGLAFTDVDVDSIKIAWESPQGQVSRVRYTSSPEDGIHELFP APDGEEDTAELOGLRPGSEYTVSVVALHDDMESQPLIGTQSTAI PAPTDLKFTQVTPTSL AQWTPPNVQLTGYRVVTPKEKTGPMKEINLAPDSSSVVVSGLMVATKYEVSVYALKDTLT SRPAQGVVTTLENVSPRRARVTDATETITISWRTKTETITGFQVDAVPANGQTPIQRTI KPDVRSYTTITGLQPGTDYKIYLYTLNDNARSSPVVIDASTAIDAPSNLRLFLATTPNSLLVS WQPPRARITGYIIKYEKPGSPPREVVPRPRPGVTEATITGLEPGTEYTIYVIALKNNQKSE PLIGRKKTGWCHDNGVNYKIGEKWDRQGENGQMMSCCLGNGKGEFKCDPHEATCYDDGKT YHVGEQWQKEYLGAICSCTCFGGQRGWRCDNCRPPGGEPSPEGTTGQSYNQYSQRYHQRTN TNVNCPIECFMPLDVQADREDSRE
--	---

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 1B.

Table 1B. Comparison of NOV1a against NOV1b.		
Protein Sequence	NOV1a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV1b	1..1951 36..1987	1370/1961 (69%) 1496/1961 (75%)

Three polymorphic variants of NOV1b have been identified and are shown in Table 41A

- 5 Further analysis of the NOV1a protein yielded the following properties shown in Table 1C.

Table 1C. Protein Sequence Properties NOV1a	
PSort analysis:	0.8800 probability located in nucleus; 0.1695 probability located in lysosome (lumen); 0.1000 probability located in mitochondrial matrix space; 0.0000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

- 10 A search of the NOV1a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 1D.

Table 1D. Geneseq Results for NOV1a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV1a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU74674	Human fibronectin protein - <i>Homo sapiens</i> , 2324 aa. [WO200187071-A1, 22- NOV-2001]	1..2320 5..2324	2320/2320 (100%) 2320/2320 (100%)	0.0

AAG68182	Fibronectin protein SEQ ID NO:98 - <i>Homo sapiens</i> , 2328 aa. [WO200177327-A1, 18-OCT-2001]	1..2320 9..2328	2320/2320 (100%) 2320/2320 (100%)	0.0
AAR92778	Human fibronectin - <i>Homo sapiens</i> , 2324 aa. [WO9604304-A1, 15-FEB-1996]	1..2320 5..2324	2318/2320 (99%) 2318/2320 (99%)	0.0
AAP70373	Human fibronectin gene product - <i>Homo sapiens</i> , 2327 aa. [EP207751-A, 07-JAN-1987]	1..2320 8..2327	2318/2320 (99%) 2318/2320 (99%)	0.0
AAM38649	Human polypeptide SEQ ID NO 1794 - <i>Homo sapiens</i> , 2355 aa. [WO200153312-A1, 26-JUL-2001]	1..2320 36..2355	2316/2320 (99%) 2317/2320 (99%)	0.0

In a BLAST search of public sequence databases, the NOV1a protein was found to have homology to the proteins shown in the BLASTP data in Table 1E.

Table 1E. Public BLASTP Results for NOV1a				
Protein Accession Number	Protein/Organism/Length	NOV1a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P02751	Fibronectin precursor (FN) (Cold-insoluble globulin) (CIG) - <i>Homo sapiens</i> (Human), 2386 aa.	1..2320 36..2386	2318/2351 (98%) 2318/2351 (98%)	0.0
FNHU	fibronectin precursor [validated] - human, 2386 aa.	1..2320 36..2386	2318/2351 (98%) 2318/2351 (98%)	0.0
E981236	FN PLASMID PFHDEL1 MATURE PROTEIN FROM PATENT WO9013653 - vectors, 2231 aa.	1..1946 5..2025	1703/2026 (84%) 1765/2026 (87%)	0.0
P07589	Fibronectin (FN) - <i>Bos taurus</i> (Bovine), 2265 aa.	1..2183 5..2213	1642/2239 (73%) 1786/2239 (79%)	0.0
P04937	Fibronectin precursor (FN) - <i>Rattus norvegicus</i> (Rat), 2477 aa.	1..2114 37..2071	1393/2128 (65%) 1584/2128 (73%)	0.0

PFam analysis predicts that the NOV1a protein contains the domains shown in Table

5 1F.

Table 1F. Domain Analysis of NOV1a			
Pfam Domain	NOV1a Match Region	Identities/ Similarities for the Matched Region	Expect Value
fn1	17..52	19/41 (46%) 35/41 (85%)	7.9e-17
fn1	62..100	21/41 (51%) 39/41 (95%)	3.2e-19
fn1	106..144	21/41 (51%) 36/41 (88%)	1.6e-17
fn1	151..190	23/41 (56%) 37/41 (90%)	4.7e-21
fn1	196..235	26/41 (63%) 38/41 (93%)	4.6e-20
fn1	273..307	14/41 (34%) 31/41 (76%)	8.1e-13
fn2	325..366	27/42 (64%) 42/42 (100%)	8e-35
fn2	385..426	26/42 (62%) 42/42 (100%)	4.3e-37
fn1	435..473	21/41 (51%) 39/41 (95%)	4.4e-20
fn1	483..520	20/41 (49%) 35/41 (85%)	2.3e-16
fn1	526..564	22/41 (54%) 37/41 (90%)	1.7e-18
fn3	573..656	28/87 (32%) 65/87 (75%)	1.7e-12
fn3	685..765	25/85 (29%) 64/85 (75%)	1.7e-14
fn3	776..854	34/84 (40%) 70/84 (83%)	1.9e-25
fn3	872..951	28/86 (33%) 63/86 (73%)	5.5e-22
fn3	962..1040	27/84 (32%) 67/84 (80%)	8.7e-21
fn3	1052..1127	26/86 (30%) 60/86 (70%)	0.0035
fn3	1139..1221	32/87 (37%) 66/87 (76%)	5.9e-19

fn3	1232..1312	27/85 (32%) 69/85 (81%)	1.8e-21
fn3	1323..1402	32/84 (38%) 68/84 (81%)	1.9e-22
fn3	1413..1495	33/86 (38%) 72/86 (84%)	4e-27
fn3	1507..1586	32/85 (38%) 69/85 (81%)	4.3e-21
fn3	1597..1676	29/86 (34%) 63/86 (73%)	2.7e-15
fn3	1687..1766	31/85 (36%) 64/85 (75%)	2.6e-20
fn3	1779..1857	30/84 (36%) 66/84 (79%)	1.6e-21
fn3	1868..1947	31/86 (36%) 69/86 (80%)	1.8e-24
fn3	2038..2115	25/87 (29%) 61/87 (70%)	5.2e-06
fn1	2140..2179	19/41 (46%) 40/41 (98%)	3.1e-20
fn1	2185..2222	21/41 (51%) 37/41 (90%)	9.4e-19
fn1	2229..2264	18/41 (44%) 36/41 (88%)	7.6e-16

Example 2.

The NOV2 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 2A.

Table 2A. NOV2 Sequence Analysis			
	SEQ ID NO: 5	1309 bp	
NOV2a, CG122589-01 DNA Sequence	GCCAGGGTTGCCTGCGGGAGCCAGGCGTCCGCTCTCCACACCTTTCACAGCCCCAGCCCTC AGAGCAAACCTCAGCCCAGCCCCAGCTCCAGCTCCAGCTCCAGCCCCGGGCCCATCAT GGCCAAGGACTTTCAAGATATCCAGCAGCTGAGCTCGGAGGAAATGACCATCCTTTCCAT CAAGGTGAGGGGCCAGGCACTCGCAGGCTGAATCCCAGGAGAGGAAATCCATTTTGAAG GGCCACCTCCTGCCAGCCCCCTGGCACAGCGTCTCTGCTCCATGGTCTGCTTCAGTCTGCT TGCCCTGAGCTTCAACATCCTGCTGCTGGTGGTCATCTGTGTGACTGGGTCCCAAAGTGAG GGTCACAGAGGTGCACAGCTGCAAGCCGAGCTGCGGAGCCTGAAGGAAGCTTTCAGCAACT TCTCCTCGAGCACCTGACGGAGGTCCAGGCAATCAGCACCCACGGAGGCAGCGTGGGTGA CAAGATCACATCCCTAGGAGCCAAGCTGGAGAAACAGCAGCAGGACCTGAAAGCAGATCAC GATGCCCTGCTCTTCCATCTGAAGCACTTCCCCGTGGACCTGCGCTTCGTGGCCTGCCAGA TGGAGCTCCTCCACAGCAACGGCTCCCAAAGGACCTGCTGCCCGTCAACTGGGTGGAGCA CCAAGGCAGCTGCTACTGTTCTCTCACTCCGGGAAGGCCTGGGCTGAGGCGGAGAAGTAC TGCCAGCTGGAGAAACGCACACCTGGTGGTCATCAACTCCTGGGAGGAGCAGAAATTCATTG TACAACACACGAACCCCTTCAATACCTGGATAGGTCTCACGGACAGTGATGGCTCTTGGAA ATGGGTGGATGGCACAGACTATAGGCACAACACAAGAAGTGGGCTGTCACTCAGCCAGAT		

	AATTGGCACGGGCACGAGCTGGGTGGAAGTGAAGACTGTGTTGAAGTCCAGCCGGATGGCC GCTGGAACGATGACTTCTGCTGCAGGTGTACCGCTGGGTGTGTGAGAAAAGCGGAATGC CACCGGCGAGGTGGCCTGACCCAGCACACCTCTGGCTAACCCATACCCACACCTGCCCA GCTCTGGCTTCTCTGTTGAGGATTTTGAGGAAAGGAAGAAACACTGAGACAGGGGTATGGG GAAGAGCTGAGCAAAGAGAGAAAGGAGGTAGTTTAAGAGTCCCTGACCCTGGAGGACTGAG ATCCACCTCCTTCTGTAATTCATTGTAATTATTATAATCGTCAGCCTCTTCAATGGCGTA GGAAAGAAGAAACAAATGCTTGAATCTC		
	ORF Start: ATG at 121		ORF Stop: TGA at 1054
	SEQ ID NO: 6	311 aa	MW at 35191.1kD
NOV2a, CG122589-01 Protein Sequence	MAKDFQDIQQLSSEENDHPPHQEGPGTRRLNPRRGNPFLKGPPLAQLRCSMVCFSL LALSFNILLVVICVTGSQSEHGRGAQLQAE LRSLKEAFSNFSSSTL TEVQAI STHGGSVG DKITSLGAKLEKQQDLKADHDALLFHLKHPVDLRFVACQMELLSNGSQRTPCPVNWVE HQGSCYWFSHSGKAWAEAEKYCLEN AHLVVINSWEEQKFIVQHTNPFNTWIGLTDSDGSW KWVDGTDYRHNYKNWAVTQPDNWHGHELGGSEDCVEVQPDGRWNDDFCLQVYRWVCEKRRN ATGEVA		
	SEQ ID NO: 7	1112 bp	
NOV2b, CG122589-02 DNA Sequence	GCCAGGGTTGCCTGCGGGAGCCAGGCGTCCGCTCTCCACACCTTTACAGCCCCAGCCCTC AGAGCAACCTCAGCCCAGCCCAGCTCCAGCTCCAGCTCCAGCCCCGGGCCCATCAT GGCCAAGGACTTTCAAGATATCCAGCAGCTGAGCTCGGAGGAAAATGACCATCCTTTCCAT CAAGGTGAGGGGCCAGGCACTCGCGGCTGAATCCAGGAGAGGAAATCCATTTTGAAG GGCCACCTCCTGCCAGCCCCTGGCACAGCGTCTCTGCTCCATGGTCTGCTTCAGTCTGCT TGCCCTGAGCTTCAACATCCTGCTGCTGGTGGTCACTGTGTGACTGGGTCCCAAAGTGCA CAGCTGCAAGCCGAGCTGCGGAGCCTGAAGGAAGCTTTCAGCAACTTCTCCTCGAGCACCC TGACGGAGGTTCAAGCAATCAGCACCCACGGAGGCAGCGTGGGTGACAAGATCACATCCCT AGGAGCCAAGCTGGAGAAACAGCAGCAGGACCTGAAAGCAGATCAGATGCCCTGCTCTTC CATCTGAAGCACTTCCCGTGGACCTGCGCTTCGTGGCCTGCCAGATGGAGCTCCTCCACA GCAACGGCTCCCAAAGGACCTGCTGCCCCGTCAACTGGGTGGAGCACCAAGGCAGCTGCTA CTGGTTCTCTCACTCCGGAAGGCTGGGCTGAGGCGGAGAAGTACTGCTGCTGGAGAAC GCACACCTGGTGGTCACTCAACTCCTGGGAGGAGCAGAAATTCATTGTACAACACACGAACC CCTTCAATACCTGGATAGGTCTCACGGACAGTGATGGCTCTTGGAATGGGTGGATGGCAC AGACTATAGGCACAACACAAGAACTGGGCTGTCACTCAGCCAGATAATTGGCACGGGCAC GAGCTGGGTGGAAGTGAAGACTGTGTTGAAGTCCAGCCGGATGGCCGCTGGAACGATGACT TCTGCCTGCAGGTGTACCGATGGGTGTGTGAGAAAAGCGGAATGCCACCGGCGAGGTGGC CTGACCCAGCACACCTCTGGCTAACCCATACCCACACCTGCCAGCTCTGGCTTCTCTG TTGAGGATTTGAG		
	ORF Start: ATG at 121		ORF Stop: TGA at 1039
	SEQ ID NO: 8	306 aa	MW at 34540.4kD
NOV2b, CG122589-02 Protein Sequence	MAKDFQDIQQLSSEENDHPPHQEGPGTRGLNPRRGNPFLKGPPLAQLRCSMVCFSL LALSFNILLVVICVTGSQSAQLQAE LRSLKEAFSNFSSSTL TEVQAI STHGGSVGDKITS LGAKLEKQQDLKADHDALLFHLKHPVDLRFVACQMELLSNGSQRTPCPVNWVEHQGSC YWFHSGKAWAEAEKYCLEN AHLVVINSWEEQKFIVQHTNPFNTWIGLTDSDGSWKWVDG TDYRHNYKNWAVTQPDNWHGHELGGSEDCVEVQPDGRWNDDFCLQVYRWVCEKRRNATGEV A		
	SEQ ID NO: 9	1055 bp	
NOV2c, CG122589-03 DNA Sequence	GCCAGGGTTGCCTGCGGGAGCCAGGCGTCCGCTCTCCACACCTTTACAGCCCCAGCCCTC AGAGCAACCTCAGCCCAGCCCAGCTCCAGCTCCAGCTCCAGCCCCGGGCCCATCAT GGCCAAGGACTTTCAAGATATCCAGCAGCTGAGCTCGGAGGAAAATGACCATCCTTTCCAT CAAGGGCCACCTCCTGCCAGCCCCTGGCACAGCGTCTCTGCTCCATGGTCTGCTTCAGTC TGCTTGCCCTGAGCTTCAACATCCTGCTGCTGGTGGTCACTGTGTGACTGGGTCCCAAAG TGCACAGCTGCAAGCCGAGCTGCGGAGCCTGAAGGAAGCTTTCAGCAACTTCTCCTCGAGC ACCTGACGGAGGTCCAGGCAATCAGCACCCACGGAGGCAGCGTGGGTGACAAGATCACAT CCCTAGGAGCCAAGCTGGAGAAACAGCAGCAGGACCTGAAAGCAGATCAGATGCCCTGCT CTTCCATCTGAAGCACTTCCCGTGGACCTGCGCTTCGTGGCCTGCCAGATGGAGCTCCTC CACAGCAACGGCTCCCAAAGGACCTGCTGCCCCGTCAACTGGGTGGAGCACAAGGCAGCT GCTACTGGTTCTCTCACTCCGGAAGGCTGGGCTGAGGCGGAGAAGTACTGCCAGCTGGA GAACGCACACCTGGTGGTCACTCAACTCCTGGGAGGAGCAGAAATTCATTGTACAACACAGC		

	AACCCTTCAATACCTGGATAGGTCTCACGGACAGTGATGGCTCTTGGAAATGGGTGGATG GCACAGACTATAGGCACAACACTACAAGAACTGGGCTGTCACTCAGCCAGATAATTGGCACGG GCACGAGCTGGGTGGAAGTGAAGACTGTGTTGAAGTCCAGCCGGATGGCCGCTGGAACGAT GACTTCTGCCTGCAGGTGTACCGTGGGTGTGTGAGAAAAGGCGGAATGCCACCGGCGAGG TGGCCTGACCCACAGCACACCTCTGGCTAACCCATACCCACACCTGCCAGCTCTGGCTTC TCTGTTGAGGATTTTGAG		
	ORF Start: ATG at 121		ORF Stop: TGA at 982
	SEQ ID NO: 10	287 aa	MW at 32550.1kD
NOV2c, CG122589-03 Protein Sequence	MAKDFQDIQQLSSEENDHPPHQGPPPAQPLAQRLCSMVCFSLALSFNILLVVICVTGSQ SAQLQAE LRSLKEAFS NFSSSTL TEVQAISTHGGSVGDKITSLGAKLEKQQDLKADHDAL LFHLKHFPVDLRFVACQME LLHNSGSRQTCCPVNWVEHQGSCYWFSHSGKAWAEAEKYCQL ENAHLVVINSWEEQKFIVQHTNPFNTWIGLTDSDGSKWVDGTDYRHNKYNWAVTQPDNWH GHELGGSEDCVEVQPDGRWNDDFCLQVYRWVCEKRRNATGEVA		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 2B.

Table 2B. Comparison of NOV2a against NOV2b and NOV2c.		
Protein Sequence	NOV2a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV2b	1..311 1..306	291/311 (93%) 291/311 (93%)
NOV2c	1..311 1..287	274/311 (88%) 274/311 (88%)

Further analysis of the NOV2a protein yielded the following properties shown in Table 2C.

Table 2C. Protein Sequence Properties NOV2a	
PSort analysis:	0.7900 probability located in plasma membrane; 0.7060 probability located in microbody (peroxisome); 0.3000 probability located in Golgi body; 0.2000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	Cleavage site between residues 3 and 4

- 5 A search of the NOV2a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 2D.

Table 2D. Geneseq Results for NOV2a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV2a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAW15246	Asialoglycoprotein receptor L-H2 - <i>Homo sapiens</i> , 287 aa. [EP773289-A2, 14-MAY- 1997]	1..311 1..287	287/311 (92%) 287/311 (92%)	e-171

AAW15252	Asialoglycoprotein receptor L-H2 cytoplasmic+extracellular domains - Chimeric <i>Homo sapiens</i> , 270 aa. [EP773289-A2, 14-MAY-1997]	1..311 1..270	270/311 (86%) 270/311 (86%)	e-159
AAW15251	Asialoglycoprotein receptor L-H2 extracellular domain - Chimeric <i>Homo sapiens</i> , 229 aa. [EP773289-A2, 14-MAY-1997]	83..311 1..229	226/229 (98%) 227/229 (98%)	e-140
AAW15245	Asialoglycoprotein receptor H1 - <i>Homo sapiens</i> , 291 aa. [EP773289-A2, 14-MAY-1997]	1..301 1..278	173/301 (57%) 214/301 (70%)	e-103
AAW15250	Asialoglycoprotein receptor H1 cytoplasmic+extracellular domains - Chimeric <i>Homo sapiens</i> , 274 aa. [EP773289-A2, 14-MAY-1997]	1..301 1..261	162/301 (53%) 200/301 (65%)	1e-95

In a BLAST search of public sequence databases, the NOV2a protein was found to have homology to the proteins shown in the BLASTP data in Table 2E.

Table 2E. Public BLASTP Results for NOV2a				
Protein Accession Number	Protein/Organism/Length	NOV2a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P07307	Asialoglycoprotein receptor 2 (Hepatic lectin H2) (ASGP-R) (ASGPR) - <i>Homo sapiens</i> (Human), 311 aa.	1..311 1..311	311/311 (100%) 311/311 (100%)	0.0
P24721	Asialoglycoprotein receptor 2 (Hepatic lectin 2) (MHL-2) (ASGP-R) (ASGPR) - <i>Mus musculus</i> (Mouse), 301 aa.	1..307 1..300	198/307 (64%) 225/307 (72%)	e-114
LNRT2	hepatic lectin 2 - rat, 301 aa.	1..307 1..300	191/307 (62%) 225/307 (73%)	e-112
P08290	Asialoglycoprotein receptor R2/3 (Hepatic lectin 2/3) (RHL-2) (ASGP-R) (ASGPR) - <i>Rattus norvegicus</i> (Rat), 301 aa.	1..307 1..300	189/307 (61%) 223/307 (72%)	e-109

AAH32130	Asialoglycoprotein receptor I - <i>Homo sapiens</i> (Human), 291 aa.	1..301 1..278	173/301 (57%) 213/301 (70%)	e-103
----------	--	------------------	--------------------------------	-------

PFam analysis predicts that the NOV2a protein contains the domains shown in Table 2F.

Table 2F. Domain Analysis of NOV2a			
Pfam Domain	NOV2a Match Region	Identities/ Similarities for the Matched Region	Expect Value
lectin_c	194..302	51/127 (40%) 99/127 (78%)	9e-50

Example 3.

The NOV3 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 3A.

Table 3A. NOV3 Sequence Analysis			
	SEQ ID NO: 11	3934 bp	
NOV3a, CG133274-01 DNA Sequence	TCCAGTAAGGAGTCGGGGTCTTCCCCAGTTTCTCAGCCAGGCGGCGGGCGGCGACTGGCAA TGTTTGGCCTCAAAAGAAACGCGGTAATCGGACTCAACCTCTACTGTGGGGGGGCCGGCTT GGGGGCGGCGAGCGGCGGCCACCCGCCGGGAGGGCGACTTTTGGCTACCGAGAAGGAG GCCTCGGCCCGGCGAGAGATAGGGGAGGGGAGGCCGCGCGGTGATTGGCGGAAGCGCCG GCGCAAGCCCCCGTCCACCTCACGCCAGACTCCCGAGGGTTCGCGCGGCCGCCGCCCAT TGGCGCCGAGGTCCCCGACGTACCCGCGACCCCCGCGAGGCTGCTTTTCTTCGCGCCCCACC CGCCGCGCGCGCCGCTTGAGGAGATGGAAGCCCCGCGCGCTGACGCCATCATGTGCCCCG AAGAGGAGCTGGACGGGTACGAGCCGGAGCCTCTCGGGAAGCGGCCGGCTGTCTGCCGCT GCTGGAGTTGGTCGGGAATCTGGTAATAACACCACTACGGACGGGTCACTACCTCGACG CCGCCGCCAGCAGAGGAGGAGGAGGACGAGTTGTACCGGCAGTCGCTGGAGATTATCTCTC GGTACCTTCGGGAGCAGGCCACCCGCGCCAAGGACACAAAGCCAATGGGCAGGTCTGGGGC CACCAGCAGGAAGCGCTGGAGACCTTACGACGGGTGGGGATGGCGTCAGCGCAACCAC GAGACGGTCTTCCAAGGCATGCTTCGAAACTGGACATCAAAAACGAAGACGATGTGAAT CGTTGTCTCGAGTGATCCATGTTTTCAGCGACGGCGTAACAACTGGGGCAGGATTGT GACTCTCATTTCTTTGGTGCTTTGTGGCTAAACACTTGAAGACCATAAACCAAGAAAGC TGCATCGAACCATTAGCAGAAAGTATCACAGACGTTCTCGTAAGGACAAAACGGGACTGGC TAGTTAAACAAAGAGGCTGGGATGGGTTTGTGGAGTTCTTCCATGTAGAGGACCTAGAAGG TGGCATCAGGAATGTGCTGCTGGCTTTTGCAGGTGTTGCTGGAGTAGGAGCTGGTTTGGCA TATCTAATAAGATAGCCTTACTGTAAGTGCAATAGTTGACTTTTAACCAACCACCACCACC ACCAAAACAGTTTATGCAGTTGGACTCCAAGCTGTAACCTCCTAGAGTTGCACCCTAGCA ACCTAGCCAGAAAAGCAAGTGGCAAGAGGATTATGGCTAACAAAGAATAAATACATGGGAAG AGTGCTCCCCATTGATTGAAGAGTCAGTGTCTGAAAGAAGCAAAGTTCAAGTTTACAGCAACA AACAACTTTGTTTGGGAAGCTATGGAGGAGGACTTTAGATTAGTGAAGATGGTAGGGT GGAAAGACTTAATTCCTTGTGAGAACAGGAAAGTGGCCAGTAGCCAGGCAAGTCATAGA ATTGATTACCCGCCGAATTCATTAATTTACTGTAGTAGTGTTAAGAGAAGCACTAAGAATG CCAGTGACCTGTGTAAGTAAGTACAGTAATAGAACTATGACTGTAAGCCTCAGTACTGTAC AAGGGAAGCTTTTCTCTCTCAATTAGCTTTCCAGTATACTTCTTAGAAAGTCCAAGTG TTCAGGACTTTTATACCTGTTATACCTTTGGCTTGGTTCCATGATCTTACTTTATTAGCCT AGTTTATACCAATAATACTTGACGGAAGGCTCAGTAATTAGTTATGAATATGGATATCCT CAATTCTAAGACAGCTTGTAAATGTATTTGTAAAAATGTATATATTTTACAGAAAGTC TATTTCTTGAACGAAGGAAGTATCGAATTTACATTAGTTTTTTTACATCCCTTTTGAAC TTTGCAACTTCCGTAATTAGGAACCTGTTTCTTACAGCTTTTCTATGCTAAACTTTGTCTCT		

119

CG133274-02 Protein Sequence	TKPMGRSGATSRKALETLLRRVGDGVQRNHETAFQGMRLRKLDIKNEDDVKSLSRVMIHVFS GVTNWGRIVTLISFGAFVAKHLKTINQESCIEPLAESITDVLVTRTKRDWLVKQRGWDGFVE FFHVEDLEGGIRNVLLAFAGVAGVGAGLAYLIR		
	SEQ ID NO: 15	667 bp	
NOV3c, 278876765 DNA Sequence	CACCGGATCCATGTTTGGCCTCAAAGAAAACGCGGTAATCGGACTCAACCTCTACTGTGGG GGGGCCGGCTTGGGGGCGGCAGCGCGCGCCACCCGCCGGGAGGGCGACTTTTGGCTA CGGAGAAGGAGGCCTCGGCCGCGCAGAGATAGGGGAGGGGAGGCCGCGCGGTGATTGG CGCCAAGGACACAAAGCCAATGGGCAGGTCTGGGGCCACCAGCAGGAAGGCGCTGGAGACC TTACGACGGGTTGGGGATGGCGTGCAGCGCAACCACGAGACGGCCTTCCAAGGCATGCTTC GGAACTGGACATCAAAAACGAAGACGATGTGAAATCGTTGTCTCGAGTGATGATCCATGT TTTCAGCGACGGCGTAACAACTGGGGCAGGATTGTGACTCTCATTCTTTTGGTGCCTTT GTGGCTAAACACTTGAAGACCATAAACCAAGAAAGCTGCATCGAACCATTAGCAGAAAGTA TCACAGACGTTCTCGTAAGGACAAAACGGGACTGGCTAGTTAAACAAAGAGGCTGGGATGG GTTTGTGGAGTTCTTCCATGTAGAGGACCTAGAAGGTGGCATCAGGAATGTGCTGCTGGCT TTTGCAGGTGTTGCTGGAGTAGGAGCTGGTTTGGCATATCTAATAAGAGTCGACGGC		
	ORF Start: at 2	ORF Stop: end of sequence	
	SEQ ID NO: 16	222 aa	MW at 23624.8kD
NOV3c, 278876765 Protein Sequence	TGSFMFLKRNAVIGLNLVYCGGAGLGAGSGGATRPGGRLATEKEASARREIGGGEAGAVIG AKDTKPMGRSGATSRKALETLLRRVGDGVQRNHETAFQGMRLRKLDIKNEDDVKSLSRVMIHV FSDGVTNWGRIVTLISFGAFVAKHLKTINQESCIEPLAESITDVLVTRTKRDWLVKQRGWDG FVEFFHVEDLEGGIRNVLLAFAGVAGVGAGLAYLIRVDG		
	SEQ ID NO: 17	610 bp	
NOV3d, 278881214 DNA Sequence	CACCGGATCGGGCTTGGGGCCGGCAGCGCGCGCCACCCGCCGGGAGGGCGACTTTTG GCTACGGAGAAGGAGGCCTCGGCCGCGCAGAGATAGGGGAGGGGAGGCCGCGCGTGA TTGGCGCCAAGGACACAAAGCCAATGGGCAGGTCTGGGGCCACCAGCAGGAAGGCGCTGGA GACCTTACGACGGGTTGGGGATGGCGTGCAGCGCAACCACGAGACGGCCTTCCAAGGCATG CTTCGGAACTGGACATCAAAAACGAAGACGATGTGAAATCGTTGTCTCGAGTGATGATCC ATGTTTTTCAGCGACGGCGTAACAACTGGGGCAGGATTGTGACTCTCATTCTTTTGGTGC CTTTGTGGCTAAACACTTGAAGACCATAAACCAAGAAAGCTGCATCGAACCATTAGCAGAA AGTATCACAGACGTTCTCGTAAGGACAAAACGGGACTGGCTAGTTAAACAAAGAGGCTGGG ATGGGTTTGTGGAGTTCTTCCATGTAGAGGACCTAGAAGGTGGCATCAGGAATGTGCTGCT GGCTTTTGCAGGTGTTGCTGGAGTAGGAGCTGGTTTGGCATATCTAATAAGAGTCGACGGC		
	ORF Start: at 2	ORF Stop: end of sequence	
	SEQ ID NO: 18	203 aa	MW at 21645.5kD
NOV3d, 278881214 Protein Sequence	TGSGLGAGSGGATRPGGRLATEKEASARREIGGGEAGAVIGAKDTKPMGRSGATSRKALE TLRRVGDGVQRNHETAFQGMRLRKLDIKNEDDVKSLSRVMIHVFS GVTNWGRIVTLISFGA FVAKHLKTINQESCIEPLAESITDVLVTRTKRDWLVKQRGWDGFVEFFHVEDLEGGIRNVLL AFAGVAGVGAGLAYLIRVDG		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 3B.

Table 3B. Comparison of NOV3a against NOV3b through NOV3d.		
Protein Sequence	NOV3a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV3b	194..350 60..216	140/157 (89%) 140/157 (89%)
NOV3c	194..350 63..219	140/157 (89%) 140/157 (89%)

NOV3d	194..350 44..200	140/157 (89%) 140/157 (89%)
-------	---------------------	--------------------------------

Further analysis of the NOV3a protein yielded the following properties shown in Table 3C.

Table 3C. Protein Sequence Properties NOV3a	
PSort analysis:	0.7300 probability located in plasma membrane; 0.6400 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in outside
SignalP analysis:	Cleavage site between residues 20 and 21

- A search of the NOV3a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
- 5 homologous proteins shown in Table 3D.

Table 3D. Geneseq Results for NOV3a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV3a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE02462	Human Mcl-1 protein - <i>Homo sapiens</i> , 350 aa. [WO200136594-A1, 25- MAY-2001]	1..350 1..350	350/350 (100%) 350/350 (100%)	0.0
AAR68814	Human mcl-1 gene product - <i>Homo sapiens</i> , 350 aa. [WO9429330-A, 22-DEC- 1994]	1..350 1..350	349/350 (99%) 349/350 (99%)	0.0
ABB57224	Mouse ischaemic condition related protein sequence SEQ ID NO:570 - <i>Mus musculus</i> , 331 aa. [WO200188188-A2, 22-NOV-2001]	1..350 1..331	266/350 (76%) 289/350 (82%)	e-144
AAE02463	Human Mcl-1s/deltaTM variant protein - <i>Homo sapiens</i> , 271 aa. [WO200136594-A1, 25- MAY-2001]	1..230 1..230	230/230 (100%) 230/230 (100%)	e-129
AAU76554	Murine Bcl-2 polypeptide - <i>Mus sp</i> , 236 aa. [WO200205835-A2, 24- JAN-2002]	193..319 66..199	45/139 (32%) 65/139 (46%)	2e-08

In a BLAST search of public sequence databases, the NOV3a protein was found to have homology to the proteins shown in the BLASTP data in Table 3E.

Table 3E. Public BLASTP Results for NOV3a				
Protein Accession Number	Protein/Organism/Length	NOV3a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
A47476	BCL2 homolog MCL1 - human, 350 aa.	1..350 1..350	350/350 (100%) 350/350 (100%)	0.0
Q9UNJ1	Myeloid cell differentiation protein (Myeloid cell leukemia protein 1) (Myeloid cell leukemia sequence 1) (BCL2-related) - <i>Homo sapiens</i> (Human), 350 aa.	1..350 1..350	349/350 (99%) 349/350 (99%)	0.0
Q07820	Induced myeloid leukemia cell differentiation protein Mcl-1 - <i>Homo sapiens</i> (Human), 350 aa.	1..350 1..350	348/350 (99%) 349/350 (99%)	0.0
Q9Z1P3	Mcl-1 protein - <i>Rattus norvegicus</i> (Rat), 330 aa.	1..350 1..330	271/350 (77%) 286/350 (81%)	e-144
P97287	EAT/MCL-1 protein (MCL1) (Myeloid cell leukemia sequence 1) - <i>Mus musculus</i> (Mouse), 331 aa.	1..350 1..331	266/350 (76%) 289/350 (82%)	e-144

PFam analysis predicts that the NOV3a protein contains the domains shown in Table 3F.

Table 3F. Domain Analysis of NOV3a			
Pfam Domain	NOV3a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Bcl-2	213..312	35/108 (32%) 100/108 (93%)	1.3e-46

Example 4.

The NOV4 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 4A.

Table 4A. NOV4 Sequence Analysis			
	SEQ ID NO: 19	1076 bp	

NOV4a, CG134430-01 DNA Sequence	TCGTGGTGCTTGGGTGGTCGCCACCAAGAAGACTTTGGTGGGGTAGTCTCGGGGCAGCTCA GCGGCCCGCTGTGCCCGTTTCTGGCCTCGCTCGCAGCTTGCACGTGAGACTCGTAGGCCG CACCCTAGGGCGAGCGTGGGGTGGCGCCGCGCGCCCTCGGGGTCTGGGCCAGCCGCA GCCTCTTCTACCGCGCCGTTGGGAGTCGCCGCGAGATGCAGCCTCCGGGCCGCCCGG GCCTATGCCCCACTAACGGGACTTCACCTTTGTCTCCTCAGCAGACCGGAAGATCTCA GTGGTTCAATAGCATCCCAGATGTCAAATTAATCTTGGTGGAGATTTATCAAAGAATC TACAGCTACTACATTTCTGAGACAAAGAGGTTATGGCTGGCTTCTGGAAGTTGAAGATGAT GATCCTGAAGATAACAAGCCACTCTTGAAGAATTGGACATTGATCTAAAGGATATTACT ACAAAATCCGATGTGTTTGTATGCCAATGCCATCACTTGGTTTAAATAGACAAGTGGTGAG AGACAATCCTGACTTTTGGGGTCTCTGGCTGTTGTTCTTTTCTTTCCATGATATCATT TATGGACAGTTTAGGGTGGTCTCATGGATTATAACCATTGGATATTTGGTTCATAACAA TTTTCTTACTGGCCAGAGTTCTTGGTGGAGAAGTTGCATATGCCCAAGTCTTGGAGTTAT AGGATATTCATTACTTCTCTCATTGTAATAGCCCTGTACTTTTGGTGGTGGATCATTT GAAGTGGTGTCTACACTTATAAAGTGAGAAGCACCAGAGGGACAGGACTTCTAGAAGTTA GAATAATATGAAGTAATCAGGAAATATCTATGCCCTACAGAAGCAGCAACCGTAAGATAAAC ATTTGTTACACTTAAGAAATTGCTGAGGTTAATACTTTGTTATAATGGATTATAATATTG ACATTCATAGTGTGACCTGGAATCTTTCACAGAAAGCTTGGGGGTGAGGACCAGGAGGT AGAATTTACAAGGCAATAAATGAAGTCTTTTAAGATC		
	ORF Start: ATG at 221		ORF Stop: TGA at 863
	SEQ ID NO: 20	214 aa	MW at 23585.1kD
NOV4a, CG134430-01 Protein Sequence	MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSIASPDVKLNLGDFIKESTATFLRQRGYG WLLEVEDDDPEDNKLLEELDIDLKDIYKIRCVLMPMPSLGFNRQVVRDNPDFWGPLAVV LFFSMISLYGQFRVSWIITWIFGSLTIFLLARVLGGEVAYQVLGVIGYSLPLIVIAPI VLLVVGSEFEVSTLIKVRSTRGTGLLEVR II		

One polymorphic variant of NOV4a has been identified and is shown in Table 41B.
Further analysis of the NOV4a protein yielded the following properties shown in Table 4B.

Table 4B. Protein Sequence Properties NOV4a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV4a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
5 homologous proteins shown in Table 4C.

Table 4C. Geneseq Results for NOV4a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV4a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB89547	Human polypeptide SEQ ID NO 1923 - <i>Homo sapiens</i> , 244 aa. [WO200190304-A2, 29-NOV-2001]	1..200 1..200	199/200 (99%) 200/200 (99%)	e-113

AAM40701	Human polypeptide SEQ ID NO 5632 - <i>Homo sapiens</i> , 316 aa. [WO200153312-A1, 26-JUL-2001]	1..200 73..272	199/200 (99%) 200/200 (99%)	e-113
AAM38915	Human polypeptide SEQ ID NO 2060 - <i>Homo sapiens</i> , 341 aa. [WO200153312-A1, 26-JUL-2001]	1..200 98..297	199/200 (99%) 200/200 (99%)	e-113
ABB11939	Human secreted protein homolog, SEQ ID NO:2309 - <i>Homo sapiens</i> , 274 aa. [WO200157188-A2, 09-AUG-2001]	1..200 31..230	199/200 (99%) 200/200 (99%)	e-113
ABG02475	Novel human diagnostic protein #2466 - <i>Homo sapiens</i> , 297 aa. [WO200175067-A2, 11-OCT-2001]	20..108 209..297	82/89 (92%) 85/89 (95%)	2e-42

In a BLAST search of public sequence databases, the NOV4a protein was found to have homology to the proteins shown in the BLASTP data in Table 4D.

Table 4D. Public BLASTP Results for NOV4a				
Protein Accession Number	Protein/Organism/Length	NOV4a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9BSR8	Similar to RIKEN cDNA 2310034L04 gene - <i>Homo sapiens</i> (Human), 244 aa.	1..200 1..200	199/200 (99%) 200/200 (99%)	e-112
Q99KZ9	Hypothetical 32.8 kDa protein - <i>Mus musculus</i> (Mouse), 289 aa.	26..200 69..245	169/177 (95%) 174/177 (97%)	2e-92
Q9CYG0	2310034L04Rik protein - <i>Mus musculus</i> (Mouse), 140 aa.	1..138 1..140	135/140 (96%) 137/140 (97%)	2e-74
Q9UIY8	Y60A3A.19 protein - <i>Caenorhabditis elegans</i> , 255 aa.	29..195 40..206	89/168 (52%) 118/168 (69%)	7e-46
Q9XTX4	T08D2.6 protein - <i>Caenorhabditis elegans</i> , 69 aa.	59..112 13..65	33/54 (61%) 40/54 (73%)	2e-11

PFam analysis predicts that the NOV4a protein contains the domains shown in Table

Table 4E. Domain Analysis of NOV4a			
Pfam Domain	NOV4a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 5.

The NOV5 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 5A.

Table 5A. NOV5 Sequence Analysis			
	SEQ ID NO: 21	1050 bp	
NOV5a, CG137677-01 DNA Sequence	TCCAGGCAACGCTGCGGCTCCGCCACGTCATGGCGCCGAGGAGAACGCGGGACAGAAC TCTGGCTGCAGGGTTTCGAGCGCGCTTCTGGCGGCGCGCTCACTGCGCTCCTTCCCCTG GCAGAGCTTAGAGGCAAAGTTAAGAGACTCATCAGATTCTGAGCTGCTGCGGGATATTTG CAGAAGACTGTGAAGCATCCCGTGTGTGAAGCACCCGCCATCAGTCAAGTATGCCGGT GCTTTCTCTCAGAACTCATCAAAAGGTCACTGCTGTCCACACGAGCCTTTGGACGAGCT GTACGAGGTGCTGGCGGAGACTCTGATGGCCAAGGAGTCCACCCAGGGCCACCGGAGCTAT TTGCTGCCCTCGGGAGGCTCGTTCACACTTCCGAGATCACAGCCATCATCTCCCATGGTA CTACAGGCTGGTCACATGGGACGCCACCCTCTACCTTGCAATGGGCCATCGAGAACCC AGCAGCCTTCACTAACAGGGGTGCTCTAGAGCTTGGCAGTGGCGCTGGCCTCACAGGCTG GCCATCTGCAAGATGTGTGCCCCAGGCATACATCTTCAAGCAGTGTACAGCCGGGTCC TCGAGCAGCTCCGAGGGAATGTCTTCTCAATGGCCTCTCATTAGAGGCAGACATCACTGC CAACTTAGACGCCCCAAGGGTGACAGTGGCCAGCTGGACTGGGACGTAGCGACAGTCCAT CAGCTCTCTGCCTTCCAGCCAGATATTGTATTGCAGCAGACGTGCTGTATTGCCAGAA CCATCGTGTCACTGGTTCGGGGTCTGCGGAGGCTGGCTGCCTGCCGGGAGCACAAGCAGGC TCCTGAGGTCTACCTGGCCTTTACCGTCCGCAACCCAGAGACGTGCCAGCTGTTCACCACC GAGCTAGGTTGGACTGGGATCAGATGGGAAGTGGAAGCTCATCATGACCAGAACTGTTTC CCTACAGAGAGCACTTGAGATGGCAATGCTGAACCTCACACTGTAGGACTCACACAGCAG TCCAACGGGCTTG		
	ORF Start: ATG at 31		ORF Stop: TAG at 1021
	SEQ ID NO: 22	330 aa	MW at 36826.8kD
NOV5a, CG137677-01 Protein Sequence	MAPEENAGTELWLQGFERRFLAARSLRSPWQSLEAKLRDSSDSELLRDIQKTVKHPVCV KHPPSVKYARCFSELIKKVS AVHTEPLDELYEVLAEETLMKESTQGHRSYLLPSGGSTL SEITAIISHGTTGLVTWDTLYLAEWAIENPAFTNRGVLELGSGAGLTGLAICKMCRPQA YIFSDCHSRVLEQLRGNVLLNGLSLEADITANLDAPRVTVQQLDWDVATVHQLSAFQPDIV IAADVLYCPEAIVSLVGLRRLAACREHKQAPEVYLAFVTRNPETCQLFTTELGWGTGIRWE VEAHDQKLFPHYREHLEMAMNLTL		

Further analysis of the NOV5a protein yielded the following properties shown in

5 Table 5B.

Table 5B. Protein Sequence Properties NOV5a	
PSort analysis:	0.7000 probability located in plasma membrane; 0.3902 probability located in microbody (peroxisome); 0.2000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV5a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 5C.

Table 5C. Geneseq Results for NOV5a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV5a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB36613	Human FLEXHT-35 protein sequence SEQ ID NO:35 - <i>Homo sapiens</i> , 330 aa. [WO200070047-A2, 23- NOV-2000]	1..330 1..330	302/330 (91%) 312/330 (94%)	e-174
ABG13115	Novel human diagnostic protein #13106 - <i>Homo sapiens</i> , 425 aa. [WO200175067-A2, 11- OCT-2001]	1..297 23..319	274/297 (92%) 284/297 (95%)	e-158
ABG09575	Novel human diagnostic protein #9566 - <i>Homo sapiens</i> , 379 aa. [WO200175067-A2, 11- OCT-2001]	1..330 1..379	259/379 (68%) 277/379 (72%)	e-134
ABG13114	Novel human diagnostic protein #13105 - <i>Homo sapiens</i> , 490 aa. [WO200175067-A2, 11- OCT-2001]	1..297 1..346	227/346 (65%) 245/346 (70%)	e-113
AAU33207	Novel human secreted protein #3698 - <i>Homo sapiens</i> , 352 aa. [WO200179449-A2, 25- OCT-2001]	33..297 8..246	209/266 (78%) 217/266 (81%)	e-108

- 5 In a BLAST search of public sequence databases, the NOV5a protein was found to have homology to the proteins shown in the BLASTP data in Table 5D.

Table 5D. Public BLASTP Results for NOV5a				
Protein Accession Number	Protein/Organism/Length	NOV5a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value

Q96G04	Similar to RIKEN cDNA 5730409G15 gene - <i>Homo sapiens</i> (Human), 330 aa.	1..330 1..330	302/330 (91%) 312/330 (94%)	e-174
Q96S85	Hypothetical 33.0 kDa protein - <i>Homo sapiens</i> (Human), 296 aa.	1..330 1..296	272/330 (82%) 282/330 (85%)	e-152
Q9CS89	5730409G15Rik protein - <i>Mus musculus</i> (Mouse), 319 aa (fragment).	1..298 1..297	214/298 (71%) 242/298 (80%)	e-117
BAC05241	CDNA FLJ40819 fis, clone TRACH2010771 - <i>Homo sapiens</i> (Human), 153 aa.	1..159 1..125	113/159 (71%) 116/159 (72%)	1e-53
Q9NVL1	CDNA FLJ10661 fis, clone NT2RP2006106 - <i>Homo sapiens</i> (Human), 165 aa.	1..114 1..87	79/114 (69%) 83/114 (72%)	4e-33

PFam analysis predicts that the NOV5a protein contains the domains shown in Table 5E.

Table 5E. Domain Analysis of NOV5a			
Pfam Domain	NOV5a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 6.

The NOV6 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 6A.

Table 6A. NOV6 Sequence Analysis			
	SEQ ID NO: 23	948 bp	
NOV6a, CG137697-01 DNA Sequence	TCCAGGCAACGCTGCGGCTCCGCCCACGTCATGGCGCCCGAGGAGAACGCGGGACAGAAC TCTGGCTGCAGGGTTTCGAGCGCCGCTTCCTGGCGGCGCGCTCACTGCGCTCCTTCCCCTG GCAGAGCTTAGAGGCAAAGTTAAGAGACTCATCAGATTCTGAGCTGCTGCGGGATATTTTG CAGAAGACTGTGAAGCATCCCGTGTGTGAAGCACCAGCCATCAGTCAAGTATGCCCGGT GCTTCTCTCAGAACTCATCAAAAGCCCTCGGGAGGCTCGTTACACTTTCGAGATCAC AGCCATCATCTCCCATGGTACTACAGGCCTGGTCACATGGGACGCCACCTCTACCTTGCA GAATGGGCCATCGAAGACCCAGCAGCCTTCACTAACAGGGGTGCTTAGAGCTTGGCAGTG GCGCTGGCCTCACAGGCCTGGCCATCTGCAAGATGTGTCGCCCCAGGCATACATCTTCAG CGACTGTCACAGCCGGGTCTCGAGCAGCTCCGAGGGAATGTCCTTCTCAATGGCCTCTCA TTAGAGGCAGACATCACTGCCAATTAGACGCCCCAAGGGTGACAGTGGCCAGCTGGACT GGGACGTAGCGACAGTCCATCAGCTCTCTGCCTTCCAGCCAGATATTGTCAATTGCAGCAGA CGTGCTGTATTGCCAGAAAGCATCGTGTCACTGGTGGGGTCTGCGGAGGCTGGCTGCC TGCCGGGAGCACAAAGCAGGCTCTGAGGTCTACCTGGCCTTACCGTCCGCAACCCAGAGA CGTGCCAGCTGTTCAACACCGAGCTAGGTTGGACTGGGATCAGATGGGAAGTGAAGCTCA TCATGACCAGAACTGTTTCCCTACAGAGAGCACTTGGAGATGGCAATGCTGAACCTCAC CTGTAGGACTCACACAGACTCCAACGGGCTTG		

	ORF Start: ATG at 31		ORF Stop: TAG at 919
	SEQ ID NO: 24	296 aa	MW at 33013.5kD
NOV6a, CG137697-01 Protein Sequence	MAPEENAGTELWLQGFERRFLAARSLRSPWQSLEAKLRDSSDSELLRDILQKTVKHPVCV KHPPSVKYARCFLSELIKPPSGGSFTLSEITAIISHGTTGLVTWDATLYLAEWAIENPAAF TNRGVLELGSGAGLTGLAICKMRPQAYIFSDCHSRVLEQLRGNVLLNGLSLEADITANLD APRVTVAQLDWDVATVHQLSAFQPDIVIAADVLYCPEAIVSLVGLRRLAACREHKQAPEV YLAFTVRNPETCOLFTTELGTGIRWEVEAHHDQKLFYPYREHLEMAMLNLT		

Further analysis of the NOV6a protein yielded the following properties shown in Table 6B.

Table 6B. Protein Sequence Properties NOV6a	
PSort analysis:	0.7000 probability located in plasma membrane; 0.4382 probability located in microbody (peroxisome); 0.2000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

- A search of the NOV6a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
- 5 homologous proteins shown in Table 6C.

Table 6C. Geneseq Results for NOV6a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV6a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB36613	Human FLEXHT-35 protein sequence SEQ ID NO:35 - <i>Homo sapiens</i> , 330 aa. [WO200070047-A2, 23- NOV-2000]	1..296 1..330	271/330 (82%) 281/330 (85%)	e-151
ABG13115	Novel human diagnostic protein #13106 - <i>Homo sapiens</i> , 425 aa. [WO200175067-A2, 11- OCT-2001]	1..263 23..319	243/297 (81%) 253/297 (84%)	e-135
ABG09575	Novel human diagnostic protein #9566 - <i>Homo sapiens</i> , 379 aa. [WO200175067-A2, 11- OCT-2001]	19..296 89..379	220/299 (73%) 233/299 (77%)	e-114

ABG13114	Novel human diagnostic protein #13105 - <i>Homo sapiens</i> , 490 aa. [WO200175067-A2, 11-OCT-2001]	19..263 89..346	188/266 (70%) 203/266 (75%)	7e-94
AAU33207	Novel human secreted protein #3698 - <i>Homo sapiens</i> , 352 aa. [WO200179449-A2, 25-OCT-2001]	33..263 8..246	183/242 (75%) 194/242 (79%)	9e-92

In a BLAST search of public sequence databases, the NOV6a protein was found to have homology to the proteins shown in the BLASTP data in Table 6D.

Table 6D. Public BLASTP Results for NOV6a				
Protein Accession Number	Protein/Organism/Length	NOV6a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96S85	Hypothetical 33.0 kDa protein - <i>Homo sapiens</i> (Human), 296 aa.	1..296 1..296	272/296 (91%) 282/296 (94%)	e-157
Q96G04	Similar to RIKEN cDNA 5730409G15 gene - <i>Homo sapiens</i> (Human), 330 aa.	1..296 1..330	271/330 (82%) 281/330 (85%)	e-151
Q9CS89	5730409G15Rik protein - <i>Mus musculus</i> (Mouse), 319 aa (fragment).	1..264 1..297	189/298 (63%) 216/298 (72%)	5e-98
BAC05241	CDNA FLJ40819 fis, clone TRACH2010771 - <i>Homo sapiens</i> (Human), 153 aa.	1..125 1..125	113/125 (90%) 116/125 (92%)	6e-59
AAH32519	Similar to hypothetical protein FLJ10661 - <i>Homo sapiens</i> (Human), 131 aa.	1..70 1..66	51/70 (72%) 58/70 (82%)	7e-20

PFam analysis predicts that the NOV6a protein contains the domains shown in Table

5 6E.

Table 6E. Domain Analysis of NOV6a			
Pfam Domain	NOV6a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 7.

The NOV7 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 7A.

Table 7A. NOV7 Sequence Analysis			
	SEQ ID NO: 25	1525 bp	
NOV7a, CG137717-01 DNA Sequence	GCGGCCGCCGAGTGAAGCAACGCGGCAACCGGAGCCCGGCGGGCAGCCGGGGAGGCCGGGA CTGAGAGGGGCGAGCCGCTGGTGTCTCCCGGCGGCAGAGGGCCGCGTCCGCCACGGGCCCG GGAGAGACGCGCTCCAGCCGCCCCAGGATGTAGGCGATCGGCGGCAGCGCTCCTGCAGGC GGCCGGCTCATCATGAAGAAGCACTCGGCCCGGGTGGCCCCGCTCAGCGCCTGCAACAGTC CGTCTGACCCCTTACCAAAGTGAAGGGGAGGAGCGCCCCGGGACTCCCCGGGCCCGGC GGAGGCCCAGGCACCGGCCGGGTGGAGGCCGGCGGGAGAGCGAGTCGCGCGTCTGGACG TGCTCCCGGGCGCAACTCAAGAAGATCTTCTGGGGCGTGGCGGTCTGCTGTGCGTGTGCT CCTCGTGGCGGGCTCCACGCAGCTCGCCAAGCTGACCTTCAGGAAGTTCGACGCGCCCTT CACCTCACGTGGTTTGCCACCAACTGGAACCTTTTATTCTTCCCGTGTACTACGTGGGG CAGTCTGCAAGTCCACAGAGAAGCAGTCTGTGAAGCAGCGATACAGGGAATGCTGTCTGAT TTTTTGGAGACAATGGCTTGACTTTGAAGGTGTTTTTACCAAGGCAGCACCCCTTTGGTGT TCTTTGGACACTCACAACTACCTGTACTTACATGCAATAAAGAAAATAAACACTACGGAT GTCTCCGTGTTGTTCTGCTGCAACAAAGCTTTGTGTTCTTGCTCTCATGGATCGTTCTCA GGGACAGATTCATGGGAGTGATTGTGGCCGCCATCCTCGCCATCGCTGGCATTGTGATGAT GACCTACGCTGATGGCTTCCACAGCCACTCCGTCTATCGGCATCGCACTGGTGGTGGCCTCA GCATCGGTTTTGTTCAAGCTCCTCCTGGGCAGTGCTAAGTTTGGAGAAGCCGCTTATTTT TGTCCATCTGGGTGTGTTAACATCCTCTTCATCACCTGCATTCCTATTATCCTCTACTT TACCAAAGTGAATACTGGAGCTCTTTGATGACATTCATGGGGAAACCTTTGTGGATTT TCAGTTCTTTATTGGCATTCAATATTGTATTAAATTTTGAATTGCGGTACATATCCCA CTCTGATGTCTCTTGAATCGTCTCAGCATACCTGTGAATGCAGTGATTGATCACTACAC CAGTCAGATCGTCTCAATGGGTCCGGGTCTATCGCCATCATCATCATCGGCCTGGGTTTT CTCCTCCTGCTCCTGCCAGAGGAGTGGGATGTCTGGTTGATCAAGCTGCTCACCCGACTCA AAGTGAGGAAGAAGGAGGAGCCTGCAGAGGGCGCTGCCGACCTGAGCTCAGGACCTCAGAG CAAGAACAAGAAGAGCCCGCCTTCCTTCGCCCGCTAACACCACTCCTCTAGAAGTCCGTGG TAATGACTGGGAGGTCTATTCTCCTGCCGGGAGGAACCTCAGTTGGGTGAAGGTGATACATCCT		
	ORF Start: ATG at 196		ORF Stop: TAA at 1438
	SEQ ID NO: 26	414 aa	MW at 45936.7kD
NOV7a, CG137717-01 Protein Sequence	MKKHSARVAPLSACNSPVLTLTKVEGEERPRDSPGPAEAQAPAGVEAGGRASRRCWTCSTRA QLKKIFWGVAVVLCVSSWAGSTQLAKLFRKFDAPFTLTWFTNWNFLFFPLYVGVHVC STEKQSVKQRYRECCRFPGDNLTLKVFFTKAAPFGVLWTLTNLYLHAIKKINTTDSVL FCCNKAFVLLSWIVLRDRFMGVIAAILAIGIVMMTYADGFHSHSVIGIALVVASASVL FKLLLSAKFGEAALFLSILGVNIFITCPIIILYFTKVEYWSFDDIPWGNLCGFSVLL LAFNIVLNFIAVTPYPTLMSLGIVLSIPVNAVIDHYTSQIVFNGVRVIAIIIGLGLLLLL LPEEWDVWLIKLLTRLKVRKKEEPAEGAADLSSGPQSKNRRARPSFAR		

Further analysis of the NOV7a protein yielded the following properties shown in

5 Table 7B.

Table 7B. Protein Sequence Properties NOV7a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4663 probability located in mitochondrial inner membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV7a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 7C.

Table 7C. Geneseq Results for NOV7a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV7a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABG16671	Novel human diagnostic protein #16662 - <i>Homo sapiens</i> , 531 aa. [WO200175067-A2, 11-OCT-2001]	5..284 168..492	160/329 (48%) 208/329 (62%)	2e-80
ABB89266	Human polypeptide SEQ ID NO 1642 - <i>Homo sapiens</i> , 134 aa. [WO200190304-A2, 29-NOV-2001]	1..134 1..134	134/134 (100%) 134/134 (100%)	1e-76
AAM36449	Peptide #10486 encoded by probe for measuring placental gene expression - <i>Homo sapiens</i> , 77 aa. [WO200157272-A2, 09-AUG-2001]	338..414 1..77	77/77 (100%) 77/77 (100%)	5e-37
AAM76340	Human bone marrow expressed probe encoded protein SEQ ID NO: 36646 - <i>Homo sapiens</i> , 77 aa. [WO200157276-A2, 09-AUG-2001]	338..414 1..77	77/77 (100%) 77/77 (100%)	5e-37
AAM63526	Human brain expressed single exon probe encoded protein SEQ ID NO: 35631 - <i>Homo sapiens</i> , 77 aa. [WO200157275-A2, 09-AUG-2001]	338..414 1..77	77/77 (100%) 77/77 (100%)	5e-37

- 5 In a BLAST search of public sequence databases, the NOV7a protein was found to have homology to the proteins shown in the BLASTP data in Table 7D.

Table 7D. Public BLASTP Results for NOV7a

Protein Accession Number	Protein/Organism/Length	NOV7a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
BAC04479	CDNA FLJ37712 fis, clone BRHIP2018369 - <i>Homo sapiens</i> (Human), 490 aa.	27..414 96..490	387/395 (97%) 387/395 (97%)	0.0
Q9JJG8	Brain cDNA, clone MNCb-0335 - <i>Mus musculus</i> (Mouse), 335 aa.	114..406 26..325	179/300 (59%) 227/300 (75%)	1e-99
Q8T0 m8	GH20388p - <i>Drosophila melanogaster</i> (Fruit fly), 578 aa.	94..379 245..536	102/295 (34%) 165/295 (55%)	7e-46
Q95XC7	Hypothetical 37.3 kDa protein - <i>Caenorhabditis elegans</i> , 339 aa.	66..368 16..326	110/320 (34%) 170/320 (52%)	5e-39
Q9VDJ2	CG15688 protein - <i>Drosophila melanogaster</i> (Fruit fly), 365 aa.	94..211 245..361	47/119 (39%) 70/119 (58%)	2e-17

PFam analysis predicts that the NOV7a protein contains the domains shown in Table 7E.

Table 7E. Domain Analysis of NOV7a			
Pfam Domain	NOV7a Match Region	Identities/ Similarities for the Matched Region	Expect Value
DUF6	78..222	24/147 (16%) 99/147 (67%)	0.053

Example 8.

The NOV8 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 8A.

Table 8A. NOV8 Sequence Analysis			
	SEQ ID NO: 27	898 bp	
NOV8a, CG137793-01 DNA Sequence	TAAGCACCCAGGAGTCCATGAAGAAGATGGCTCCTGCCATGGAATCCCCTACTCTACTGTGT GTAGCCTTACTGTTCTTCGCTCCAGATGGCGTGTAGCAGTCCCTCAGAAACCTAAGGTCT CCTTGAACCCCTCCATGGAATAGAATATTTAAAGGAGAGAATGTGACTCTTACATGTAATGG GAACAATTTCTTTGAAGTCAGTTCACCAAATGGTTCCACAATGGCAGCCTTTCAGAAGAG ACAAATTCAGTTTGAATATTGTGAATGCCAAATTTGAAGACAGTGGAGAATACAAATGTC AGCACCAACAAGTTAATGAGAGTGAACCTGTGTACCTGGAAGTCTTCAGTGACTGGCTGCT CCTTCAGGCCTCTGCTGAGGTGGTGTGAGGGGCCAGCCCCCTCTCCTCAGGTGCCATGGT TGGAGGAAGTGGGATGTGTACAAGGTGATCTATTATAAGGATGGTGAAGCTCTCAAGTACT GGTATGAGAACCACAACATCTCCATTACAATGCCACAGTTGAAGACAGTGGAACTACTA CTGTACGGGCAAAGTGTGGCAGCTGGACTATGAGTCTGAGCCCCCTCAACATTACTGTAATA AAAGCTCCGCGTGAGAAGTACTGGCTACAATTTTTATCCCATTGTTGGTGGTGATTCTGT		

	TTGCTGTGGACACAGGATTATTATCTCAACCCAGCAGCAGGTCACATTTCTCTGAAGAT TAAGAGAACCAGGAAAGGCTTCAGACTTCTGAACCCACATCCTAAGCCAAACCCAAAAAC AACTGATATAATTACTCAAGAAATATTTGCAACATTAGTTTTTTTCCAGCATCAGCAATTG CTACTCAATTGTCAAACACAGCTTGCAATAAAGGGCGATTCCAG		
	ORF Start: ATG at 26		ORF Stop: TGA at 797
	SEQ ID NO: 28	257 aa	MW at 29595.6kD
NOV8a, CG137793-01 Protein Sequence	MAPAMESPTLLCVALLFFAPDGVLA VPQKP KVS LNPPWNRI FKGENVT LTCNGN NFFEVSS TKWFHNGSLSEETNSSLNIVNAKFEDSGEYKCHQHQVNESEPVYLEVFSDWLLQASAEVV MEGQPLFLRCHGWRNWDVYKVIYYKDGEALKYWYENHNISITNATVEDSGTYICTGKVWQL DYSEPLNITVIKAPREKYWLQFFIPLLVLFAVD TGLFISTQQQVTFLLKIKRTRKGFR LLNPHPKPNPKNN		
	SEQ ID NO: 29	757 bp	
NOV8b, CG137793-02 DNA Sequence	TAAGCACCAGGAGTCCATGAAGAAGATGGCTCCTGCCATGGAATCCCTACTCTACTGTGT GTAGCCTTACTGTTCTTCGCTCCAGATGGCGTGTTAGCAGTCCCTCAGAAACCTAAGGTCT CCTTGAACCCTCCATGGAATAGAATATTTAAAGGAGAGAATGTGACTCTTACATGTAATGG GAACAATTTCTTTGAAGTCAGTTCACCAAATGGTTCCACAATGGCAGCCTTTTCAAGAGAG ACAAATTCAGTTTGAATATTGTGAATGCCAAATTTGAAGACAGTGGAGAATACAAATGCC ATGGTTGGAGGAAGTGGGATGTGTACAAGGTGATCTATTATAAGGATGGTGAAGCTCTCAA GTACTGGTATGAGAACCACAACATCTCCATTACAAATGCCACAGTGAAGACAGTGGGAACC TACTACTGTACGGGCAAAGTGTGGCAGCTGGACTATGAGTCTGAGCCCCTCAACATTACTG TAATAAAAGTCCGCGTGAGAAGTACTGGCTACAATTTTTATCCCATTTGTTGGTGGTGAT TCTGTTTGCTGTGGACACAGGATTATTTATCTCAACTCAGCAGCAGGTCACATTTCTCTTG AAGATTAAGAGAACCAGGAAAGGCTTCAGACTTCTGAACCCACATCCTAAGCCAAACCCCA AAAACAACTGATATAATTACTCAAGAAATATTTGCAACATTAGTTTTTTTCCAGCATCAGC AATTGCTACTCAATTGTCAAACACA		
	ORF Start: ATG at 26		ORF Stop: TGA at 680
	SEQ ID NO: 30	218 aa	MW at 25079.5kD
NOV8b, CG137793-02 Protein Sequence	MAPAMESPTLLCVALLFFAPDGVLA VPQKP KVS LNPPWNRI FKGENVT LTCNGN NFFEVSS TKWFHNGSLSEETNSSLNIVNAKFEDSGEYKCHGWRNWDVYKVIYYKDGEALKYWYENHNIS ITNATVEDSGTYICTGKVWQLDYSEPLNITVIKAPREKYWLQFFIPLLVLFAVD TGLFISTQQQVTFLLKIKRTRKGFRLLNPHPKPNPKNN		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 8B.

Table 8B. Comparison of NOV8a against NOV8b.		
Protein Sequence	NOV8a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV8b	1..246 1..207	207/246 (84%) 207/246 (84%)

Twenty polymorphic variants of NOV8b have been identified and are shown in Table 41C.

- 5 Further analysis of the NOV8a protein yielded the following properties shown in Table 8C.

Table 8C. Protein Sequence Properties NOV8a

PSort analysis:	0.4600 probability located in plasma membrane; 0.1594 probability located in microbody (peroxisome); 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 26 and 27

A search of the NOV8a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 8D.

Table 8D. Geneseq Results for NOV8a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV8a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB31584	Amino acid sequence of a human Fc epsilon receptor alpha-chain - <i>Homo sapiens</i> , 257 aa. [WO200104310-A1, 18-JAN-2001]	1..257 1..257	257/257 (100%) 257/257 (100%)	e-155
AAB74667	Human immunoglobulin E receptor I alpha subunit protein - <i>Homo sapiens</i> , 257 aa. [WO200111010-A2, 15-FEB-2001]	1..257 1..257	257/257 (100%) 257/257 (100%)	e-155
AA Y96230	Human Fc receptor, FcepsilonRIa - <i>Homo sapiens</i> , 260 aa. [EP1006183-A1, 07-JUN-2000]	1..257 4..260	257/257 (100%) 257/257 (100%)	e-155
AAW61190	The alpha chain of a Fc epsilon receptor - <i>Homo sapiens</i> , 257 aa. [WO9823964-A1, 04-JUN-1998]	1..257 1..257	257/257 (100%) 257/257 (100%)	e-155
AAW24066	Alpha subunit of human high affinity receptor for IgE (human FcERI) - <i>Homo sapiens</i> , 257 aa. [US5639660-A, 17-JUN-1997]	1..257 1..257	257/257 (100%) 257/257 (100%)	e-155

In a BLAST search of public sequence databases, the NOV8a protein was found to have homology to the proteins shown in the BLASTP data in Table 8E.

Table 8E. Public BLASTP Results for NOV8a

Protein Accession Number	Protein/Organism/Length	NOV8a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P12319	High affinity immunoglobulin epsilon receptor alpha-subunit precursor (FcERI) (IgE Fc receptor, alpha-subunit) (Fc-epsilon RI-alpha) - <i>Homo sapiens</i> (Human), 257 aa.	1..257 1..257	257/257 (100%) 257/257 (100%)	e-154
AAH15195	Fc IgE, high affinity I, receptor for, alpha polypeptide - <i>Homo sapiens</i> (Human), 257 aa.	1..257 1..257	256/257 (99%) 256/257 (99%)	e-154
CAC28464	Sequence 4 from Patent WO0104310 - <i>Homo sapiens</i> (Human), 232 aa (fragment).	26..257 1..232	232/232 (100%) 232/232 (100%)	e-139
CAC28471	Sequence 26 from Patent WO0104310 - Cloning vector pINT1, 660 aa.	1..197 1..197	197/197 (100%) 197/197 (100%)	e-117
CAC28468	Sequence 17 from Patent WO0104310 - Cloning vector pINT1, 756 aa (fragment).	1..197 1..197	197/197 (100%) 197/197 (100%)	e-117

PFam analysis predicts that the NOV8a protein contains the domains shown in Table 8F.

Table 8F. Domain Analysis of NOV8a			
Pfam Domain	NOV8a Match Region	Identities/ Similarities for the Matched Region	Expect Value
ig	44..95	19/54 (35%) 37/54 (69%)	1.4e-10
ig	125..178	14/56 (25%) 37/56 (66%)	0.00018

Example 9.

The NOV9 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 9A.

Table 9A. NOV9 Sequence Analysis

	SEQ ID NO: 31	4330 bp
NOV9a, CG137873-01 DNA Sequence	TCTAGGAGCCAGCCCCACCCTTAGAAAAGATGTTTTCCATGAGGATCGTCTGCCTGGTCCCT AAGTGTGGTGGGCACAGCATGGACTGCAGATAGTGGTGAAGGTGACTTTCTAGCTGAAGGA GGAGGCGTGGTGGCCCAAGGGTTGTGGAAGACATCAATCTGCCTGCAAAGATTCAAGCT GGCCCTTCTGCTCTGATGAAGACTGGAACACAAATGCCCTTCTGGCTGCAGGATGAAAGG GTTGATTGATGAAGTCAATCAAGATTTTACAAACAGAATAAAAGCTCAAAAATTCACATA TTTGAATATCAGAAGACAATAAGGATTCTCATTGTTGACCACTAATATAATGGAATTT TGAGAGGCGATTTTTCTCAGCCAATAACCGTGATAATACCTACAACCGAGTGTGAGAGGA TCTGAGAAGCAGAATTGAAGTCTGAAGCGCAAAGTTCATAGAAAAAGTACAGCATATCCAG CTTCTGCAGAAAAATGTTAGAGCTCAGTTGGTTGATATGAAACGACTGGAGGTGGACATTG ATATTAAGATCCGATCTTGTCGAGGGTCATGCAGTAGGGCTTTAGCTCGTGAAGTAGATCT GAAGGACTATGAAGATCAGCAGAAGCAACTTGAACAGGTCAATGGCCAAAGACTTACTTCCC TCTAGAGATAGGCAACACTTACCCTGATAAAAATGAAACCAAGTCCAGACTTGGTTCCCG GAAATTTTAAAGAGCCAGCTTCAGAAGGTACCCCCAGAGTGGAAAGGCATTAACAGACATGCC GCAGATGAGAATGGAGTTAGAGAGACCTGGTGGAAATGAGATTACTCGAGGAGGCTCCACC TCTTATGGAACCGGATCAGAGACGGAAGCCCCAGGAACCTAGCAGTGTGGAAGCTGGA ACTCTGGGAGCTCTGGACCTGGAAGTACTGGAACCGAAACCTGGGAGCTCTGGGACTGG AGGGACTGCAACCTGGAAACCTGGGAGCTCTGGACCTGGAAGTACTGGAAGCTGGAACCTCT GGGAGCTCTGGAACCTGGAAGTACTGGAACCAAAACCTGGGAGCCCTAGACCTGGTAGTA CCGGAACCTGGAATCTGGCAGCTCTGAACGCGGAAGTGTGGGCACTGGACCTCTGAGAG CTCTGTATCTGGTAGTACTGGACAATGGCACTCTGAATCTGGAAGTTTATAGGCCAGATAGC CCAGGCTCTGGGAACGCGAGGCCCTAAACACCCAGACTGGGGCACATTTGAAGAGGTGTGAG GAAATGTAAGTCCAGGGACAAGGAGAGAGTACCACACAGAAAAACTGGTCACTTCTAAAGG AGATAAGAGCTCAGGACTGGTAAAGAGAAGGTACCTCTGGTAGCACAAACCACACGCGT CGTTCATGCTCTAAACCGTTACTAAGACTGTTATTGGTCTGATGGTCACAAAGAAGTTA CCAAAGAAGTGGTGACTCCGAAGATGGTCTGACTGTCCCGAGGCAATGGATTTAGGCAC ATTGTCTGGCATAGGTACTCTGGATGGGTTCGCCATAGGCACCTGATGAAGCTGCCTTC TTCGACACTGCCTCACTGGAAAAACATTCCAGGTTTCTTCTCACCCTATTAGGAGGT TTGTCACTGAGACTGACTCTAGGGCTCAGAATCTGGCATCTTCAAAATACAAAGGAATC CAGTTCTCATCACCTGGGATAGCTGAATCCCTTCCCGTGGTAAATCTTCAAGTTACAGC AAACAATTACTAGTAGCAGGATTACAACAGAGGAGACTCCACATTTGAAAGCAAGAGCT ATAAAATGGCAGATGAGGCCGGAAGTGAAGCCGATCATGAAGGAACACATAGCACCAAGAG AGGCCATGCTAAATCTCGCCCTGTGAGAGTATCCACACTTCTCCTTTGGGGAGCCCTTCC CTGTCCCTTAGACTAAGTTAAATATTTCTGCACAGTGTCCCATGGCCCCCTTGCAATTTCC TTCTTAATCTCTGTTACACGTCATTGAACTACACTTTTTTGGTCTGTTTTTGTGCTAGA CTGTAAGTTCCTTGGGGGAGGGCCTTTGTCTGTCTCATCTCTGATTCCCAATGCCTAA CAGTACAGAGCCATGACTCAATAAATACATGTTAAATGGATGAATGAATCCCTCTGAAACT CTATTTGAGCTTATTTAGTCAAATCTTTCACTATTCAAAGTGTGTGCTATTAGAAATTGTC ACCCAAGTGATTAAATCACATTTTAGTATGTGTCTCAGTTGACATTTAGGTCAAGCTAAAT ACAAGTTGTGTTAGTATTAAGTGATGCTTAGCTACCTGTACTGGTTACTTGTCTATTAGTTT GTGCAAGTAAATTCCAAATACATTTGAGGAAAATCCCTTTGCAATTTGTAGGTATAAAT AACCGCTTATTTGCATAAGTTCTATCCCCTGTAAGTGCATCCTTTCCCTATGGAGGGAAG GAAAGGAGGAAGAAAGAAAGGAAGGGAAGAAAGAAACAGTATTTGCCTTATTTAATCTGAGCCG TGCCTATCTTTGTAAAGTTAAATGAGAATAACTTCTTCCAACCAAGCTTAATTTTTTTTTTA GACTGTGATGATGTCTCCAAACACATCCTTCAGGTACCCAAAGTGGCATTTCATATCA AGCTACCGGGATCCAGTAAGATTTTTTCTGTTTATTGCGATCAAGAGACCAGTTTGGGAGG ATGGCTTTTGATCCAGCAAGAATGGATGGATCACTGAATTTTAAACCGGACTGGCAAGAC TACAAGAGAGGTTTCGGCAGCCTGAATGACGAGGGGGAAGGAGAATTCTGGCTAGGCAATG ACTACCTCCACTTACTAACCACAAAGGGGCTCTGTTCTTAGGGTTGAATTAGAGGACTGGGC TGGGAATGAAGCTTATGAGAATATCACTTCCGGGTAGGCTCTGAGGCTGAAGGCTATGCC CTCCAAGTCTCCTCTATGAAGGCACTGCGGGTGATGCTCTGATTGAGGGTTCCGTAGAGG AAGGGGAGAGTACACCTCTCACAACAACATGCAGTTCAGCACCTTTGACAGGGATGCAGA CCAGTGGGAAGAGAAGTGTGAGAAGTCTATGGGGGAGGCTGGTGGTATAATAACTGCCAA GCAGCCAATCTCAATGGAATCTACTACCTGGGGCTCCTATGACCAAGGAATAACAGTC CTTATGAGATTGAGAATGGAGTGGTCTGGGTTTCCCTTAGAGGGGAGATTTATCCCTCAG GGCTGTTCGATGAAAATTAGGCCCTTGTGACCCAATAGGCTGAAGAAGTGGGAATGGGA GCACTCTGTCTTTTGTGCTAGAGAAGTGGAGAGAAAATACAAAGGTAAAGCAGTTGAGAT TCTCTACAACCTAAAAAATCCTAGGTGCTATTTTCTTATCCTTTGTACTGTAGCTAAATG TACCTGAGACATATTAGTCTTTGAAAAATAAAGTTATGTAAGGTTTTTTTATCTTTAAT	

	AGCTCTGTGGGTTTAAACATTTTGTAAAGATATACCAAGGGCCATTGAGTACATCAGGAA AGTGGCAGACAGAAGCTTCTCTGCAACCTTGAAGACTATTGGTTTGAGAAGTCTCTTC CCATACCACCCAAAATCATAATGCCATTGGAAAGCAAAAAGTTGTTTATCCATTTGATTT GAATTGTTTTAAGCCAATATTTAAGGTAAACTCACTGAATCTAACCATAGCTGACCTTT GTAGTAGAATTTCAACTTATAATTACAATGCACAATTTATAATTACAATATGATTTATG TCTTTTGCTATGGAGCAAATCCAGGAAGGCAAGAGAAACATTCTTTCTAAATATAAATGA AAATCTATCCTTTAAACTCTTCCACTAGACGTTGTAATGCACACTATTTTTTTCCCAAGG AGTAACCAATTTCTTTCTAAAACACATTTAAAATTTTAAACTATTATGAATATTAAAA AAGACATAATTCACACATTAATAAACAATCTCCCAAGTATTGATTTAACTTCATTTTTCTA ATAATCATAAACTATATTCTGTGACATGCTAATTATTATTAAATGAAGTCGTTAGTTCGA AAGCCTCTCACTAAGTATGATCTATGCTATATTCAAAATCAACCCATTACTTTGGTCAA TATTTGATCTAAGTTGCATCTTTAATCCTGGTGGTCTTGCCTTCTGATTTTTAATTTGTAT CCTTTCTATTAAAGATATATTGTCATTTTCTCTGAATATGTATTAAAATATCCCAAGC		
	ORF Start: ATG at 30		ORF Stop: TAG at 1962
	SEQ ID NO: 32	644 aa	MW at 69756.0kD
NOV9a, CG137873-01 Protein Sequence	MFSMRIVCLVLSVVGTAWTADSGEGDFLAEGGGVVRGPRVVERHQSACKDSWDWPCSDWDWN YKCPSGCRMKGLIDEVNQDFTNRINKLNSLFEYQKNNKDSHSLTNNIMEILRGDFSSANN RDNTYNRVSEDLRSRIEVLKRVIEKVQHIQLLQKNVRAQLVDMKRLEVDIDIKIRSCRGS CSRALAREVDLKDIEDQKQLEQVIAKDLLPSRDRQHLPLIKMKPVPLVPGNFKSQLQKV PPEWKALTDMPQMRMELERPGNEITRGGSTSYGTGSETESPRNPSSAGSWNSGSSGPGST GNRNPSSSGTGATATWKPGSSGPGSTGWSNSGSSGTGTCNQNPGSPRPGSTGTWNPSSSE RGSAGHWTSESSVSGSTGQWHSSESGFRPDSPGSGNARPNPNPDWGTFFEVSGNVSPGTRE YHTEKLVTSKGDKEKLTGKEKVTSGSTTTTTRRSCSKTVTKTVIGPDGHKEVTKEVVTSEDG SDCPEAMDGLTSLGIGTLDGFRHRHPDEAAFPDASTGKTFPGFFSPMLGEFVSETESRGS ESGI FTNTKESSSHHPGIAEFPSRKSSSYSKQFTSSTSYNRGDSTFESKSYKMADEAGSE ADHEGTHSTKRGHAKSRPVRGIHTSPLGKPSLSP		
	SEQ ID NO: 33	1515 bp	
NOV9b, CG137873-03 DNA Sequence	AATCCTTTCTTTAGCTGGAGTGTCTCAGGAGCCAGCCCCACCCTTAGAAAAGATGTTTT CCATGAGGATCGTCTGCCTGGTCTAAGTGTGGTGGGCACAGCATGGACTGCAGATAGTGG TGAAGGTGACTTTCTAGCTGAAGGAGGAGCGTGGTGGCCCAAGGGTGTGGAAGACAT CAATCTGCCTGCAAAGATTGAGCTGGCCCTTCTGCTCTGATGAAGACTGGAATACAAAT GCCCTTCTGGCTGCAGGATGAAAGGGTTGATTGATGAAGTCAATCAAGATTTTACAAACAG AATAAATAAGCTCAAAAATTCATATTGAAATATCAGAAGAACAATAAGGATTTCTATTCTG TTGACCACTAATATAATGAAATTTTGAGAGGCGATTTTCTCTAGCCAATAACCGTGATA ATACCTACAACCGAGTGTGAGGATCTGAGAAGCAGAATTGAAGTCTGAAGCGCAAAGT CATAGAAAAGTACAGCATATCCAGCTTCTGCAAAAAATGTTAGAGCTCAGTTGGTTGAT ATGAAACGACTGGAGGTGGACATTGATATTAAGATCCGATCTTGTGCGAGGTCATGCAGTA GGGCTTTAGTCTGTGAAGTAGATCTGAAGGACTATGAAGATCAGCAGAAGCAACTGAACA GGTCATTGCGCAAAGACTTACTTCCCTCTAGAGATAGGCAACACTTACCAGTATCAAAATG AAACCACTTCCAGACTTGGTTCCCGGAAATTTAAGAGCCAGCTTCAGAAGGTACCCCGAG AGTGGAAGGCATTAACAGACATGCCGAGATGAGAATGGAGTTAGAGAGACCTGGTGGAAA TGAGATTACTCGAGGAGGCTCCACCTCTTATGGAACCGGATCAGAGACGGAAGCCCCAGG AACCCTAGCAGTGTGGAAGCTGGAACCTTGGGAGCTCTGGACCTGGAAGTACTGGAAGCT GGAAGCTGGAAGTACTGGAACCAAAACCCCTGGGAGCCCTAGACCTGGTAGTACCAGAAC TGGAATCCTGGCAGCTCTGAACGCGAAGTGTGGGCACTGGACCTCTGAGAGCTCTGTAT CTGGTAGTACTGGACAATGGCACTCTGAATCTGGAAGTTTATAGGCCAGATAGCCAGGCTC TGGAACGCGAGGCTTAAACCCAGACTGGGGCACATTTGAAGAGGTGTGAGGAAATGTA AGTCCAGGGACAAGAGAGAGTACACACAGAAAACCTGGTCTTCTACAAGAGATAAGAGCTC GGACTGGTAAGAGAGGTCACTCTGGTACACACACACGCGTGTCTCTCTAAACGTACTAG ACGTATGGCCGATGTCCAGAGTACAGAATGGAACCAATGTCACTCCAGAAGATAGAATTT AGATTAATTAAGGTCCAAGCCGAATGCTAACTCATAAATGTTACCTAAAAATAGAACTGA TAATCAATTACATAATAATAAAGATAAAGATAAAAAAAGAATAAAAAAAA		
	ORF Start: ATG at 55		ORF Stop: TAA at 1219
	SEQ ID NO: 34	388 aa	MW at 43094.6kD
NOV9b, CG137873-03	MFSMRIVCLVLSVVGTAWTADSGEGDFLAEGGGVVRGPRVVERHQSACKDSWDWPCSDWDWN YKCPSGCRMKGLIDEVNQDFTNRINKLNSLFEYQKNNKDSHSLTNNIMEILRGDFSSANN RDNTYNRVSEDLRSRIEVLKRVIEKVQHIQLLQKNVRAQLVDMKRLEVDIDIKIRSCRGS		

Protein Sequence	CSRALAREVDLKDYEQQKQLEQVIAKDLLPSRDRQHLPLIKMKPVPDLVPGNFKSQLQKV PPEWKALDMPQMRMELERPGGNEITRGGSTSYGTGSETESPRNPSSAGSWSNGSSGPGST GSWKLEVLTKTLGALDLVVPFGILALNAEVLGTGPLRALYLVLDNGTLNLEVLGQIA QALGTRGLTTQTGAHLKRCQEM		
	SEQ ID NO: 35	1734 bp	
NOV9c, CG137873-02 DNA Sequence	AATCCTTTCTTTTCAGCTGGAGTGTCTCAGGAGCCAGCCCCACCTTAGAAAAAGATGTTTT CCATGAGGATCGTCTGCCTGGTCTAAGTGTGGTGGGCACAGCATGGACTGCAGATAGTGG TGAAGGTGACTTTCTAGCTGAAGGAGGAGCGTGCCTGGCCCAAGGGTTGTGAAAGACAT CAATCTGCCTGCAAAGATTTCAGACTGGCCCTTCTGCTCTGATGAAGACTGGAATACAAAT GCCCTTCTGGCTGCAGGATGAAAGGGTTGATTGATGAAGTCAATCAAGATTTTACAAACAG AATAAATAAGCTCAAAAATTCATTATTTGAATATCAGAAGAACAATAAGGATTCTCATTCTG TTGACCACTAATATAATGAAATTTTGAGAGGCGATTTTCTCAGCCAATAACCGTGATA ATACCTACAACCGAGTGTGAGAGGATCTGAGAAGCAGAATTGAAGTCTGAAGCGCAAAGT CATAGAAAAAGTACAGCATATCCAGCTTCTGCAAAAAAATGTTAGAGCTCAGTTGGTTGAT ATGAAACGACTGGAGGTGGACATTGATATTAAGATCCGATCTTGTGAGGGTCTATGCAGTA GGGCTTTAGCTCGTGAAGTAGATCTGAAGGACTATGAAGATCAGCAGAAGCAACTTGAACA GGTCATTGCCCAAAGACTTACTTCCCTCTAGAGATAGGCAACACTTACCACTGATCAAAATG AAACCACTTCCAGACTTGGTTCCTGGGAAATTTAAGAGCCAGCTTCAAGAGTACCCCCAG AGTGGAAAGGCATTAAACAGACATGCCGAGATGAGAATGGAGTTAGAGAGACTTGGTGGAAA TGAGATTACTCGAGGAGGCTCCACTTCTTATGGAACCGGATCAGAGACGGAAAGCCCAAGG AACCTAGCAGTGTGGAAGCTGGAAGTCTGGGAGCTCTGGACCTGGAAGTACTGGAAGCT GGAAGTCTGGGAGCTCTGGAAGTGAAGTACTGGAACCAAAACCTGGGAGCCCTAGACC TGGTAGTACCGGAACCTGGAATCCTGGCAGCTCTGAACGCGGAAGTGTGGGCACTGGACC TCTGAGAGCTCTGTATCTGGTAGTACTGGACAATGGCACTCTGAATCTGGAAGTTTATAGGC CAGATAGCCCAGGCTCTGGGAACGCGAGGCTTAACAACCCAGACTGGGGCTCAGAATCTGG CATCTTCACAAATACAAAGGAATCCAGTTCTCATCACCTGGGATAGCTGAATTCCTTCC CGTGGTAAATCTTCAAGTTACAGCAAACAATTTACTAGTAGCAGAGTTACAACAGAGGAG ACTCCACATTTGAAAGCAAGAGCTATAAAATGGCAGATGAGGCCGGAAGTGAAGCCGATCA TGAAGGAACACATAGCACCAAGAGAGGCCATGCTAAATCTCGCCCTGTGAGGATATCCAC ACTTCTCTTTGGGGAAGCCTTCCCTGTCCCCCTAGACTAAGTTAAATATTTCTGCACAGT GTTCCCATGGCCCTTGCAATTCCTTCTTAAGTCTCTGTTACACGTCATTGAACTACACT TTTTTGGTCTGTTTTGTGCTAGACTGTAAGTTCCCTGGGGGCGAGGCTTGTCTGTCTC ATCTCTGTATTCCTAAATGCCTAACAGTACAGGCCCATGACTCAATAAATACATGTTAAAT GGATGAATGAATTCCTCTGAACTCT		
	ORF Start: ATG at 55		ORF Stop: TAG at 1498
	SEQ ID NO: 36	481 aa	MW at 52648.5kD
NOV9c, CG137873-02 Protein Sequence	MFSMRIVCLVLSVVGTAWTADSGEGDFLAEGGGVVRGPRVVERHQSAKSDWPFCSDEEDWN YKCPSGCRMKGLIDEVNDFTNRINKLKNLSLFEYQKNNKDSHSLTTNIMEILRGDFSSANN RDNTYNRVSEDLRSRIEVLKRKVIKQVHIQLLQKNVRAQLVDMKRLEVDIDIKIRSCRGS CSRALAREVDLKDYEQQKQLEQVIAKDLLPSRDRQHLPLIKMKPVPDLVPGNFKSQLQKV PPEWKALDMPQMRMELERPGGNEITRGGSTSYGTGSETESPRNPSSAGSWSNGSSGPGST GSWNSGSGTGSTGNQNPGRPGSTGTWNPSSERGSAGHWTSESSVSGSTGQWHSESGS FRPDSPGSGNARPNPDWGSSEGIFTNTKESSSHHPGIAEFPSRGKSSSYSKQFTSSTS RGDSTFESKSYKMADEAGSEADHEGTHSTKRGHAKSRPVRGIHTSPLGKPSLSP		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 9B.

Table 9B. Comparison of NOV9a against NOV9b and NOV9c.		
Protein Sequence	NOV9a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV9b	1..289 1..289	260/289 (89%) 260/289 (89%)

NOV9c	1..412	318/412 (77%)
	1..386	319/412 (77%)

Further analysis of the NOV9a protein yielded the following properties shown in Table 9C.

Table 9C. Protein Sequence Properties NOV9a	
PSort analysis:	0.5087 probability located in outside; 0.1900 probability located in lysosome (lumen); 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 20 and 21

- A search of the NOV9a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
- 5 homologous proteins shown in Table 9D.

Table 9D. Geneseq Results for NOV9a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV9a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAR82244	Human fibrinogen A-alpha chain protein - <i>Homo sapiens</i> , 644 aa. [WO9523868-A1, 08-SEP-1995]	1..644 1..644	643/644 (99%) 643/644 (99%)	0.0
AAR60020	Fibronectin - <i>Homo sapiens</i> , 643 aa. [WO9416085-A, 21-JUL-1994]	1..644 1..643	641/644 (99%) 641/644 (99%)	0.0
AAY82891	AlphaE subunit of human fibrinogen - <i>Homo sapiens</i> , 847 aa. [WO200009562-A1, 24-FEB-2000]	20..641 1..626	615/626 (98%) 616/626 (98%)	0.0
AAR60019	Tissue-binding hybrid protein - <i>Homo sapiens</i> , 1336 aa. [WO9416085-A, 21-JUL-1994]	210..644 910..1336	416/435 (95%) 417/435 (95%)	0.0
AAB54135	Human pancreatic cancer antigen protein sequence SEQ ID NO:587 - <i>Homo sapiens</i> , 360 aa. [WO200055320-A1, 21-SEP-2000]	1..307 22..328	301/307 (98%) 301/307 (98%)	e-176

In a BLAST search of public sequence databases, the NOV9a protein was found to have homology to the proteins shown in the BLASTP data in Table 9E.

Table 9E. Public BLASTP Results for NOV9a				
Protein Accession Number	Protein/Organism/Length	NOV9a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
FGHUA	fibrinogen alpha chain precursor, short splice form [validated] - human, 644 aa.	1..644 1..644	644/644 (100%) 644/644 (100%)	0.0
P02671	Fibrinogen alpha/alpha-E chain precursor [Contains: Fibrinopeptide A] - <i>Homo sapiens</i> (Human), 866 aa.	1..641 1..645	634/645 (98%) 635/645 (98%)	0.0
P02672	Fibrinogen alpha chain [Contains: Fibrinopeptide A] - <i>Bos taurus</i> (Bovine), 596 aa (fragment).	20..644 4..596	375/633 (59%) 442/633 (69%)	0.0
Q99K47	Fibrinogen A alpha polypeptide - <i>Mus musculus</i> (Mouse), 557 aa.	1..634 1..557	371/637 (58%) 436/637 (68%)	0.0
P06399	Fibrinogen alpha/alpha-E chain precursor - <i>Rattus norvegicus</i> (Rat), 782 aa.	1..626 1..544	359/629 (57%) 428/629 (67%)	0.0

PFam analysis predicts that the NOV9a protein contains the domains shown in Table 9F.

Table 9F. Domain Analysis of NOV9a			
Pfam Domain	NOV9a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 10.

- 5 The NOV10 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 10A.

Table 10A. NOV10 Sequence Analysis			
	SEQ ID NO: 37	730 bp	

NOV10a, CG137882-01 DNA Sequence	ATGCGAACACAAGTATATGAGGGGTTGTGTAAAAATTATTTTCTCTTGCTGTACTACAA AGAGATAGAATCAAACGCTTTTTTCGACATACTGGTTTTCTTTCTGTTTTCTTCTC TTTCTTCTATTTCTTGTGGATATTATGGCTAATAACACAACAAGTTTAGGGAGTCCATGG CCAGAAAACTTTTGGGAGGACCTTATCATGTCTTCACTGTATCCATGGCAATCGGGCTG GTACTTGGAGGATTATTTGGGCTGTGTTTATTGTCTGTCTCGAAGAAGAAGAGCCAGT GCTCCCATCTCACAGTGGAGTTCAAGCAGGAGATCTAGGTCTTCTTACACCCACGGCCTC AACAGAACTGGATTTTACCGCCACAGTGGCTGTGAACGTCGAAGCAACCTCAGCCTGGCC AGTCTCACCTTCCAGCGACAAGCTTCCCTGGAACAAGCAAATTCCTTTCCAAGAAAATCA AGTTTCAGAGCTTCTACTTTCCATCCCTTTCTGCAATGTCCACCACTTCTGTGGAACCT GAGAGTCAGCTGGTGACTCTCCCTTCTTCCAATATCTCTCCACCATCAGCACTTCCCACT AGTCTGAGCCGCTCTGACTACTGGTCCAGTAACAGTCTTCGAGTGGGCCTTTCAACACCG CCCCACCTGCCTATGAGTCCATCATCAAGGCATTCCAGATTCTGAGTAGGGTGGCTTT TTGGTTTTTG		
	ORF Start: ATG at 1		ORF Stop: TGA at 706
	SEQ ID NO: 38	235 aa	MW at 26592.1kD
NOV10a, CG137882-01 Protein Sequence	MRTQVYEGLCKNYFSLAVLQDRIKLLFFDILVFLSVFLLFLLFLVDIMANNTSLGSPW PENFWEDLIMSFTVSMAGLVLGGFIWAVFICLSRRRRRASAPISQWSSRRSRSSYTHGL NRTGFYRHSGCERRSNLSLASLTFRQASLEQANSFPRKSSFRASTFHPFLQCPPLPVET ESQLVTLPSNISPTISTSHSLRPDYWSSNSLRVGLSTPPPPAYESI IKAFPDS		
	SEQ ID NO: 39	630 bp	
NOV10b, CG137882-02 DNA Sequence	ATGCGAACACAAGTATATGAGGGGTTGTGTAAAAATTATTTTCTCTTGCTGTACTACAA AGAGATAGAATCAAACGCTTTTTTCGACATACTGGTTTTCTTTCTGTTTTCTTCTC TTTCTTCTATTTCTTGTGGATATTATGGCTAATAACACAACAAGTTTAGGGAGTCCATGG CCAGAAAACTTTTGGGAGGACCTTATCATGTCTTCACTGTATCCATGGCAATCGGGCTG GTTCTTGGAGGATTATTTGGGCTGTGTTTATTGTCTGTCTCGAAGAAGAAGAGCCAGT GCTCCCATCTCACAGTGGAGTTCAAGCAGGAGATCTAGGTCTTCTTACACCCACGGCCTC AACAGAACTGGATTTTACCGCCACAGTGGCTGTGAACGTCGAAGCAACCTCAGCCTGGCC AGTCTCACCTTCCAGCGACAAGCTTCCCTGGAACAAGCAAATTCCTTTCCAATATCTCTC CCACCATCAGCACTTCCCACTGCTGAGCCGCTCTGACTACTGGTCCAGTAACAGTCTTC GAGTGGGCCTTTCAACACCGCCCCACCTGCCTATGAGTCCATCATCAAGGCATTCCCACT ATTCTGAGTAGGGTGGCTTTTGGTTTTTG		
	ORF Start: ATG at 1		ORF Stop: TGA at 505
	SEQ ID NO: 40	168 aa	MW at 19141.9kD
NOV10b, CG137882-02 Protein Sequence	MRTQVYEGLCKNYFSLAVLQDRIKLLFFDILVFLSVFLLFLLFLVDIMANNTSLGSPW PENFWEDLIMSFTVSMAGLVLGGFIWAVFICLSRRRRRASAPISQWSSRRSRSSYTHGL NRTGFYRHSGCERRSNLSLASLTFRQASLEQANSFPISLPPSALPTV		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 10B.

Table 10B. Comparison of NOV10a against NOV10b.		
Protein Sequence	NOV10a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV10b	1..157 1..157	125/157 (79%) 125/157 (79%)

Further analysis of the NOV10a protein yielded the following properties shown in Table 10C.

Table 10C. Protein Sequence Properties NOV10a

PSort analysis:	0.6000 probability located in nucleus; 0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	Cleavage site between residues 51 and 52

A search of the NOV10a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 10D.

Table 10D. Geneseq Results for NOV10a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV10a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AA Y59671	Secreted protein 108-006-5-0-C2-FL - <i>Homo sapiens</i> , 187 aa. [WO9940189-A2, 12-AUG-1999]	49..235 1..187	187/187 (100%) 187/187 (100%)	e-107
AAE01707	Human gene 5 encoded secreted protein HHBCS39, SEQ ID NO:119 - <i>Homo sapiens</i> , 166 aa. [WO200134767-A2, 17-MAY-2001]	70..235 1..166	166/166 (100%) 166/166 (100%)	1e-92
AAE01676	Human gene 5 encoded secreted protein HHBCS39, SEQ ID NO:88 - <i>Homo sapiens</i> , 166 aa. [WO200134767-A2, 17-MAY-2001]	70..235 1..166	166/166 (100%) 166/166 (100%)	1e-92
AA Y65073	Human 5' EST related polypeptide SEQ ID NO:1234 - <i>Homo sapiens</i> , 59 aa. [WO9953051-A2, 21-OCT-1999]	1..59 1..59	56/59 (94%) 56/59 (94%)	5e-24
AAG01373	Human secreted protein, SEQ ID NO: 5454 - <i>Homo sapiens</i> , 136 aa. [EP1033401-A2, 06-SEP-2000]	49..184 1..136	49/137 (35%) 57/137 (40%)	7e-11

In a BLAST search of public sequence databases, the NOV10a protein was found to have homology to the proteins shown in the BLASTP data in Table 10E.

Table 10E. Public BLASTP Results for NOV10a

Protein Accession Number	Protein/Organism/Length	NOV10a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
AAM88866	MTLC - <i>Homo sapiens</i> (Human), 235 aa.	1..235 1..235	235/235 (100%) 235/235 (100%)	e-134
Q9H763	CDNA: FLJ21269 fis, clone COL01745 - <i>Homo sapiens</i> (Human), 235 aa.	1..235 1..235	234/235 (99%) 235/235 (99%)	e-133
CAD39158	Hypothetical protein - <i>Homo sapiens</i> (Human), 204 aa (fragment).	32..235 1..204	204/204 (100%) 204/204 (100%)	e-115
Q8TBE8	Similar to RIKEN cDNA 1110020B04 gene - <i>Homo sapiens</i> (Human), 187 aa.	49..235 1..187	186/187 (99%) 186/187 (99%)	e-105
Q8R411	MT-MC1 - <i>Mus musculus</i> (Mouse), 188 aa.	49..235 1..188	160/188 (85%) 173/188 (91%)	4e-90

PFam analysis predicts that the NOV10a protein contains the domains shown in Table 10F.

Table 10F. Domain Analysis of NOV10a			
Pfam Domain	NOV10a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 11.

The NOV11 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 11A.

Table 11A. NOV11 Sequence Analysis			
	SEQ ID NO: 41	957 bp	
NOV11a, CG137910-01 DNA Sequence	CATCATGCTATGGAAAAATGGAAGAATTTGTTTGTAAAGGTATGGGAAGGTCGGTGGCGA GTGATCCCTCATGATGTACTACCAGACTGGCTCAAGGATAATGACTTCCTCTTGCATGGA CACCGCCTCCTATGCCTTCTTCCGGGCCTGTTTAAAGAGCATTTTCAGAATACACACA GAAACAGGCAACATTTGGACACATCTCTTAGGTTGTGTATTCTTCTGTGCCTGGGGATC TTTTATATGTTTCGCCCCAAATATCTCCTTTGTGGCCCCTCTGCAAGAGAAGGTGGTCTTT GGATTATTTTCTTAGGAGCATTCTCTGCCTTCTTTTTCATGGCTCTCCACACAGTC TACTGCCACTCAGAGGGGGTCTCTCGGCTCTTCTCTAAACTGGATTACTCTGGTATTGCT CTTCTGATTATGGGAAGTTTGTTCCTTGGCTTTATTATTCTTCTACTGTAATCCACAA CCTTGCTTCATCTACTTGATTGTCATCTGTGTGCTGGGCATTGCAGCCATTATAGTCTCC CAGTGGGACATGTTGCCACCCCTCAGTATCGGGGAGTAAGAGCAGGAGTGTTTTTGGGC CTAGGCCTGAGTGAATCATTCCTACCTTGCATATGTCATCTCGGAGGGGTCTCCTTAGG GCCGCCACCATAGGGCAGATAGGCTGGTTGATGCTGATGGCCAGCCTCTACATCACAGGA GCTGCCCTGTATGCTGCCCGGATCCCGAACGCTTTTCCCTGGCAAATGTGACATCTGG		

	TTTCACTCTCATCAGCTGTTTCATATCTTGTGGTTGCTGGAGCTTTTGTCACTTCCAT GGTGTCTCAAACCTCCAGGAGTTTCGTTTCATGATCGGCGGGGGCTGCAGTGAAGAGGAT GCACTGTGATACCTACCAGTCTCCAGGGACTATGACCCTAAACCAGGGCCTGCGGCA		
	ORF Start: ATG at 10		ORF Stop: TGA at 907
	SEQ ID NO: 42	299 aa	MW at 34157.9kD
NOV11a, CG137910-01 Protein Sequence	MEKMEEFVCKVWEGRWVIPHDLVLPDWLKDNDFLHGHRRPPMPSFRACFKSIFRIHTETG NIWTHLLGCVFFLCGLIFYMFRPNISFVAPLQEKVVFGFLGAILCLSFSWLFHTVYCH SEGVSRFLSKLDYSGIALLIMGSFVPWLYSFYCNPPCFIYLIVICVLGIAAIIVSQWD MFATPQYRGVRAGVFLGLGLSGIIPTLHYVISEGFLRAATIGQIGWMLMASLYITGAAL YAARIPEFFPGKCDIWFHSHQLFHFVVGAFVHFHGVSNLQEFRFMIGGGCSEEDAL		

Further analysis of the NOV11a protein yielded the following properties shown in Table 11B.

Table 11B. Protein Sequence Properties NOV11a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.3000 probability located in microbody (peroxisome)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV11a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
5 homologous proteins shown in Table 11C.

Table 11C. Geneseq Results for NOV11a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV11a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM79290	Human protein SEQ ID NO 1952 - <i>Homo sapiens</i> , 258 aa. [WO200157190-A2, 09-AUG-2001]	42..299 1..258	256/258 (99%) 257/258 (99%)	e-154
ABB89913	Human polypeptide SEQ ID NO 2289 - <i>Homo sapiens</i> , 375 aa. [WO200190304-A2, 29-NOV-2001]	1..299 77..375	238/299 (79%) 269/299 (89%)	e-149
AAB74699	Human membrane associated protein MEMAP-5 - <i>Homo sapiens</i> , 375 aa. [WO200112662-A2, 22-FEB-2001]	1..299 77..375	238/299 (79%) 269/299 (89%)	e-149

AAM79634	Human protein SEQ ID NO 3280 - <i>Homo sapiens</i> , 379 aa. [WO200157190-A2, 09-AUG-2001]	1..299 81..379	238/299 (79%) 269/299 (89%)	e-149
AAM78650	Human protein SEQ ID NO 1312 - <i>Homo sapiens</i> , 375 aa. [WO200157190-A2, 09-AUG-2001]	1..299 77..375	238/299 (79%) 269/299 (89%)	e-149

In a BLAST search of public sequence databases, the NOV11a protein was found to have homology to the proteins shown in the BLASTP data in Table 11D.

Table 11D. Public BLASTP Results for NOV11a				
Protein Accession Number	Protein/Organism/Length	NOV11a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9H737	CDNA: FLJ21432 fis, clone COL04219 - <i>Homo sapiens</i> (Human), 258 aa.	42..299 1..258	256/258 (99%) 257/258 (99%)	e-153
Q91VH1	Hypothetical 42.4 kDa protein - <i>Mus musculus</i> (Mouse), 375 aa.	1..299 77..375	238/299 (79%) 269/299 (89%)	e-149
Q96A54	Similar to CGI-45 protein (Hypothetical 42.6 kDa protein) - <i>Homo sapiens</i> (Human), 375 aa.	1..299 77..375	238/299 (79%) 269/299 (89%)	e-149
Q9Y360	CGI-45 protein - <i>Homo sapiens</i> (Human), 370 aa.	1..292 77..368	236/292 (80%) 264/292 (89%)	e-147
Q9CZA0	2810031L11Rik protein - <i>Mus musculus</i> (Mouse), 352 aa.	1..276 77..352	211/276 (76%) 236/276 (85%)	e-126

PFam analysis predicts that the NOV11a protein contains the domains shown in

5 Table 11E.

Table 11E. Domain Analysis of NOV11a			
Pfam Domain	NOV11a Match Region	Identities/ Similarities for the Matched Region	Expect Value
UPF0073	43..280	126/287 (44%) 220/287 (77%)	3.5e-125

Example 12.

The NOV12 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 12A.

Table 12A. NOV12 Sequence Analysis			
	SEQ ID NO: 43	714 bp	
NOV12a, CG138013-01 DNA Sequence	TAACCCAGAACATCTGGCACCTCTAACCCAGACATCTGGCACCTCTAACCCAGACAT CTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCTGCTCTGGGGAGGGAG AGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCGTGACGGTGCAGGAA GGCCTGTGTGTCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATTTACCCT GGCCCACTAGTTCATGGCTACTGGTTCCGGGAAGGGGCAATACAGACCAGGATGCTCCA GTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTC CTTGGGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGAT GCGGGGAGATACTTCTTTCTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCAC CGGCTCTCTGTGAATGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCAGGCACC CTGGAGTCCGGCTGCCCCCAAGAATCTGACCCACTCCTCAGTGGGGGAAGGAGAGCTCCAG TATGCATCCCTCAGCTTCCAGATGGTGAAGCCTTGGGACTCACGGGGACAGGAGGCCACT GACACCGAGTACTCGGAGATCAAGATCCACAGATGAGAAACTGCAGAGACTCAC		
	ORF Start: ATG at 82		ORF Stop: TGA at 694
	SEQ ID NO: 44	204 aa	MW at 23190.0kD
NOV12a, CG138013-01 Protein Sequence	MLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFYPSHGWYIPGPVVHGY WFREGANTDQDAPVATNNPARAVWEBTRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFR MEKGSIKWNYKHHRLSVNVLTALHTRPNILIPGTLESGCPQNLTHSSVGEGLQYASLSFQ MVKPWDSRGQEATDTEYSEIKIHR		

Further analysis of the NOV12a protein yielded the following properties shown in

5 Table 12B.

Table 12B. Protein Sequence Properties NOV12a	
PSort analysis:	0.4170 probability located in lysosome (lumen); 0.3700 probability located in outside; 0.2303 probability located in microbody (peroxisome); 0.1000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	Cleavage site between residues 18 and 19

A search of the NOV12a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 12C.

Table 12C. Geneseq Results for NOV12a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV12a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value

AAM49113	Human dendritic cell membrane protein Siglec-9 - <i>Homo sapiens</i> , 463 aa. [JP2001352977-A, 25-DEC-2001]	1..165 1..165	164/165 (99%) 164/165 (99%)	5e-97
AAU87079	Sialic acid-binding Ig-related lectin, Siglec-BMS-L5a - <i>Homo sapiens</i> , 463 aa. [WO200208257-A2, 31-JAN-2002]	1..165 1..165	164/165 (99%) 164/165 (99%)	5e-97
AAB29186	OB binding protein like protein #1 - <i>Homo sapiens</i> , 444 aa. [WO200053747-A1, 14-SEP-2000]	1..165 31..195	164/165 (99%) 164/165 (99%)	5e-97
AAB66137	Protein of the invention #49 - Unidentified, 463 aa. [WO200078961-A1, 28-DEC-2000]	1..165 1..165	164/165 (99%) 164/165 (99%)	5e-97
AAB87568	Human PRO1302 - <i>Homo sapiens</i> , 463 aa. [WO200116318-A2, 08-MAR-2001]	1..165 1..165	164/165 (99%) 164/165 (99%)	5e-97

In a BLAST search of public sequence databases, the NOV12a protein was found to have homology to the proteins shown in the BLASTP data in Table 12D.

Table 12D. Public BLASTP Results for NOV12a				
Protein Accession Number	Protein/Organism/Length	NOV12a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
AAF87223	Sialic acid-binding immunoglobulin-like lectin-9 - <i>Homo sapiens</i> (Human), 463 aa.	1..165 1..165	164/165 (99%) 164/165 (99%)	1e-96
Q9Y336	OB binding protein-like protein (Sialic acid-binding lectin) - <i>Homo sapiens</i> (Human), 463 aa.	1..165 1..165	164/165 (99%) 164/165 (99%)	1e-96
Q9BY19	FOAP-9 - <i>Homo sapiens</i> (Human), 463 aa.	1..165 1..165	163/165 (98%) 164/165 (98%)	4e-96
Q9Y286	QA79 membrane protein, allelic variant AIRM-1B precursor - <i>Homo sapiens</i> (Human), 467 aa.	1..165 2..169	132/169 (78%) 138/169 (81%)	3e-68

Q9Y502	QA79 membrane protein, splice product AIRM-2 precursor - <i>Homo sapiens</i> (Human), 374 aa.	1..140 2..144	109/144 (75%) 115/144 (79%)	6e-55
--------	---	------------------	--------------------------------	-------

Pfam analysis predicts that the NOV12a protein contains the domains shown in Table 12E.

Table 12E. Domain Analysis of NOV12a			
Pfam Domain	NOV12a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 13.

The NOV13 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 13A.

Table 13A. NOV13 Sequence Analysis			
	SEQ ID NO: 45	1240 bp	
NOV13a, CG138074-01 DNA Sequence	GGTCACTACGTGCTCTGGCCCTCCACCTCCTGGGTCAAAGCACCCAAGGCAGGGCTGAG GAGGTGGGCCAGAACCGGTTGGCGGGCAGGTGGGCACATTGTCAGAGATGGCTATCAGC CTCCTCTCCCTTCTTTCTCTGTTGCAGCTGCCTTGGGTGCGGTGCAGGGTCCAGAGCC ATTCGGACCTGTTACTACTGGGCCTCACAGGTGGCCTGACTCTGTTACTGCTCCTGACA CTGCTGGCCCTTGGCCGGTACTCAGGGCTGTTGGCTGGGTGGCAGTGAGTGCCGGCTCA CCCCCATCCGGCTCACCCCCATCCCTACAAGTCCATGTGGAGCCCTATGGTGAGACT GGGTGGCTTTTTCACCAGAGCTGCAGCATCTCCCCAAGCTCTGCTCCATCGCTGTCCAC GTTCTCTTGGCAAGTGCTGTTGTGTCTGGGCAGCCTCCTGAGTGAAGGTGAGGAATCT CCCTCCCTGAGCTCATCCACATCTACCAGAAATTGACTTCAAGGCATTCTCCTTCAG GCACCCAGCCACGTGGTGACAGCCACCTTCCCCACACCACCATGCTGTCCATCTGGGTG GCTGCCTGCCATATCCATTCTGCCTTGGACACCTACATCAAGGGAACGACATGATGAGT GACACGAGTTCTGGAAGCTTGGAGGTGAGCCCTGGTAGCCGGGAGACTTCATTTGCTACC GTGTACCTGGGGAGAGCGGCCGCGCTGGGAGGATGGTGACACCTGCAGTGAGTGCAGC TGCAGAGAGTCAGGTGCCAGCGGCTCCTCTTTTGGAGAGCTGGACCTGGAGGGTGAGGGG CCCTTGGAGGAACACGGCTGGACCTGAGACTGAGCCCCCTGGGGGCTACCAAGTGGCCC TGAGAGCCAGTACCCTGAGAAGGGCAAGGAGTAACCCATGACCAGCCCCCTCTGCGGG GCAGGGCTGCGGAACCGAGCAGACTCTCCAGCCATCTTCTCCTTCTTCTGGGGGCGAGG GGTCCCAGGGGACGTAACCTCCCCCTGCTCTAGGCCTCTTGTGAAGCCTTCTCCTCACTG TCCTTTAGGCTCCCAGGGCCAAAGCAGCCAAAGACTGTATCCTGCACCAGCCCTGTGGGC CGACACTCCTGTTGTATCTCTTTTTCAGACTGTCACTGGAGCTTCCAGGACCCAGAATAA AGCCAATGACTTACTTGTTCAAAAAAAAAAAAAAAAAAG		
	ORF Start: ATG at 109		ORF Stop: TGA at 901
	SEQ ID NO: 46	264 aa	MW at 28038.5kD
NOV13a, CG138074-01 Protein Sequence	MAISLLSLLSLQLPWGAVQGSRAISDLLLLGLTGGLTLLLLLTLAFAGYSGLLAGVAV SAGSPPIRLTPHPYKFHVEPYGETGWLHFQSCSISPKLCSIAVHVPLGKCCCVLGSLLSE GEESPSPELIHIYQKDFKAFSFPQAPSHVVTATFPYTTMLSIWVAACHIHSALDITYIKGT DMMSDTSSGSLEVSPGSRFETSFATVSPGESGRGWEDGDTCECSCRESGASGSSFEELDL EGEGPLEEPRLDPETEPLGATKWP		

Further analysis of the NOV13a protein yielded the following properties shown in Table 13B.

Table 13B. Protein Sequence Properties NOV13a	
PSort analysis:	0.4600 probability located in plasma membrane; 0.1197 probability located in microbody (peroxisome); 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 50 and 51

A search of the NOV13a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
5 homologous proteins shown in Table 13C.

Table 13C. Geneseq Results for NOV13a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV13a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE06730	Human CASB765 protein - <i>Homo sapiens</i> , 311 aa. [WO200157077-A1, 09-AUG-2001]	25..264 1..311	208/318 (65%) 215/318 (67%)	e-100
AAU81960	Human PRO536 - <i>Homo sapiens</i> , 313 aa. [WO200109327-A2, 08-FEB-2001]	25..263 1..301	174/302 (57%) 187/302 (61%)	8e-79
AAB65173	Human PRO536 (UNQ337) protein sequence SEQ ID NO:97 - <i>Homo sapiens</i> , 313 aa. [WO200073454-A1, 07-DEC-2000]	25..263 1..301	174/302 (57%) 187/302 (61%)	8e-79
AAB94830	Human protein sequence SEQ ID NO:15991 - <i>Homo sapiens</i> , 313 aa. [EP1074617-A2, 07-FEB-2001]	25..263 1..301	174/302 (57%) 187/302 (61%)	8e-79
AAU12370	Human PRO536 polypeptide sequence - <i>Homo sapiens</i> , 313 aa. [WO200140466-A2, 07-JUN-2001]	25..263 1..301	174/302 (57%) 187/302 (61%)	8e-79

In a BLAST search of public sequence databases, the NOV13a protein was found to have homology to the proteins shown in the BLASTP data in Table 13D.

Table 13D. Public BLASTP Results for NOV13a				
Protein Accession Number	Protein/Organism/Length	NOV13a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q99LS5	Similar to putative secreted protein (Unknown) (Protein for MGC:7091) - <i>Mus musculus</i> (Mouse), 309 aa.	27..259 3..296	173/294 (58%) 188/294 (63%)	2e-81
Q9D7D9	Adult male tongue cDNA, RIKEN full-length enriched library, clone:2310012P03, full insert sequence - <i>Mus musculus</i> (Mouse), 309 aa.	27..259 3..296	172/294 (58%) 187/294 (63%)	1e-80
Q9Y6I9	Putative secreted protein ZSIG11 precursor - <i>Homo sapiens</i> (Human), 313 aa.	25..263 1..301	174/302 (57%) 187/302 (61%)	2e-78
CAC25002	Sequence 46 from Patent WO0100806 precursor - <i>Homo sapiens</i> (Human), 312 aa.	25..263 1..300	173/302 (57%) 186/302 (61%)	2e-76
Q9UKD7	Hypothetical 9.7 kDa protein - <i>Homo sapiens</i> (Human), 93 aa.	183..263 1..81	67/81 (82%) 69/81 (84%)	4e-30

PFam analysis predicts that the NOV13a protein contains the domains shown in Table 13E.

Table 13E. Domain Analysis of NOV13a			
Pfam Domain	NOV13a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 14.

The NOV14 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 14A.

Table 14A. NOV14 Sequence Analysis			
	SEQ ID NO: 47	843 bp	
NOV14a, CG138573-01 DNA Sequence	GGGGTGGAGTGGGGTGTTCATTTCATCAAGTGTGCAGCATGGGTCTCTCTGTAGCAGGCC ATGGCATGCTGGTGGCCGCTCCTGCTAGAGCTGTGGACAGTCATGCCACCTGGGCTGGG GACGAGCTGCTCAACATCTGCATGAATGCCAAACACCACAAGAGAGTGCCAGCCAGAA GACAAGCTCTATGAGGAGTGCATCCCTGGAAGGACAATGCCTGCTGCACCCTCACGACA AGCTGGGAAGCCCATCTGGATGTATCCCACTCTACAACCTCAGCCTGTTCACTGTGGA		

	CTGCTGATGCCTGGCTGTGCGAAGCACTTCATCCAGGCTATCTGCTTCTATGAGTGCTCC CCAAACCTGGGGCCCTGGATCCAGCCAGTGGCCCCGAGTGGGCAGGGAGAGCGAGTTGTG AATGTGCCGCTGTGCCAGGAGGACTGTGAGGAGTGGTGGGAAGACTGTGCGATGTCTTAC ACATGCAAATCCAACCTGGCGTGGTGGCTGGGACTGGAGTCAGGGGAAGAACCGCTGCCCC AAAGGGCCCCAGTGCCTCCCTTCTCCCATTAATTCCTCCACCCAGCTGACCTGTGTGAG AAGACTTGGAGCAATTCCTTCAAAGCCAGCCCTGAGCGACGGAACAGTGGGCGGTGTCTC CAGAAGTGGTTTGAGCCTGCTCAGGGCAACCCCAATGTGGCCGTGGCCCGCTCTTCGCC AGCTCTGCCCCATCCTGGGAAGTGTCTTACACCATCATGGTCTGCTCCCTGTTCTCTGCCG TTCCTTTCCTGAGAGCCCTTCTTCTCCCACTCACATTCTGTCATGTCCACCAACTGTGGG TCA		
	ORF Start: ATG at 61		ORF Stop: TGA at 790
	SEQ ID NO: 48	243 aa	MW at 27942.7kD
NOV14a, CG138573-01 Protein Sequence	MACWWPLLELWTVMPWAGDELLNICMNAKHHKRVSPEDKLYEECI PWKDNACCTLT SWEAHLDVSPLYNPSLFHCGLLMPGCRKHFIQAICFYECSPNLGPWIQPVAPSGQGERVV NVPLQEDCEEWEDCRMSYTKSNWRGGWDWSQGNRCPKGAQCLPFSHYFPTPADLCE KTWSNSFKASPERRNSGRCLQKWFEPAQGNPNVAVARLFASSAPSWELSYTIMVCSLFLP FLS		

Further analysis of the NOV14a protein yielded the following properties shown in Table 14B.

Table 14B. Protein Sequence Properties NOV14a	
PSort analysis:	0.7480 probability located in microbody (peroxisome); 0.4420 probability located in mitochondrial matrix space; 0.1282 probability located in mitochondrial inner membrane; 0.1282 probability located in mitochondrial intermembrane space
SignalP analysis:	Cleavage site between residues 20 and 21

A search of the NOV14a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
5 homologous proteins shown in Table 14C.

Table 14C. Geneseq Results for NOV14a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV14a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE09454	Human sbg72825FOLATEa protein - <i>Homo sapiens</i> , 250 aa. [WO200160850-A1, 23-AUG-2001]	1..243 1..250	243/250 (97%) 243/250 (97%)	e-156
AAB50286	Human folate receptor II protein SEQ ID NO: 6 - <i>Homo sapiens</i> , 255 aa. [WO200071754-A1, 30-NOV-2000]	4..222 5..226	130/222 (58%) 158/222 (70%)	8e-82

ABG19167	Novel human diagnostic protein #19158 - <i>Homo sapiens</i> , 248 aa. [WO200175067-A2, 11-OCT-2001]	19..222 29..235	120/207 (57%) 144/207 (68%)	7e-70
ABG04155	Novel human diagnostic protein #4146 - <i>Homo sapiens</i> , 206 aa. [WO200175067-A2, 11-OCT-2001]	46..242 1..204	101/205 (49%) 128/205 (62%)	5e-54
ABG19166	Novel human diagnostic protein #19157 - <i>Homo sapiens</i> , 187 aa. [WO200175067-A2, 11-OCT-2001]	19..153 27..176	66/151 (43%) 81/151 (52%)	9e-30

In a BLAST search of public sequence databases, the NOV14a protein was found to have homology to the proteins shown in the BLASTP data in Table 14D.

Table 14D. Public BLASTP Results for NOV14a				
Protein Accession Number	Protein/Organism/Length	NOV14a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9EQF4	Folate receptor 3 (Folate receptor 4) (Delta) - <i>Mus musculus</i> (Mouse), 244 aa.	1..241 1..242	166/242 (68%) 191/242 (78%)	e-104
P15328	Folate receptor alpha precursor (FR-alpha) (Folate receptor 1) (Folate receptor, adult) (Adult folate-binding protein) (FBP) (Ovarian tumor- associated antigen MOv18) (KB cells FBP) - <i>Homo sapiens</i> (Human), 257 aa.	7..242 10..255	140/246 (56%) 169/246 (67%)	1e-84
Q9XSH1	Membrane-bound folate binding protein - <i>Sus scrofa</i> (Pig), 249 aa.	7..239 8..247	138/240 (57%) 167/240 (69%)	4e-84
P41439	Folate receptor gamma precursor (FR-gamma) (Folate receptor 3) - <i>Homo sapiens</i> (Human), 243 aa.	19..222 27..230	129/204 (63%) 152/204 (74%)	5e-82

P35846	Folate receptor alpha precursor (FR-alpha) (Folate receptor 1) (Folate-binding protein 1) - <i>Mus musculus</i> (Mouse), 255 aa.	7..242 10..251	135/242 (55%) 168/242 (68%)	7e-82
--------	--	-------------------	--------------------------------	-------

PFam analysis predicts that the NOV14a protein contains the domains shown in Table 14E.

Table 14E. Domain Analysis of NOV14a			
Pfam Domain	NOV14a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Folate_rec	4..238	133/243 (55%) 181/243 (74%)	4e-110

Example 15.

- 5 The NOV15 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 15A.

Table 15A. NOV15 Sequence Analysis			
	SEQ ID NO: 49	1885 bp	
NOV15a, CG138606-01 DNA Sequence	TCCTCAAATACAATGCTTCAAAAAACGCTGCTGATCTTGATCTCTTTTTCAGTAGTAACC TGGATGATTTTTATAATTTCTCAGAACTTCACAAAGCTTTGGTCTGCTCTAACTTATCC ATCTCTGTCCATTACTGGAACAACCTCCGCAAAGTCCTTATTCCTAAAAACATCACTGATA CCATTAAAGCCACTAACAGAGACTGAACTCAGAATAAAGGAAATCATAGAGAACTAGAT CAGCAGATCCCACCCAGACCTTTACCCATGTGAACACCACCAGTGCCACACACAGC ACAGCCACCATCCTCAACCTCGAGATACATACTGCAGGGGAGACCAGCTGGACATCCTA CTGGAGGTGAGGGACCCTTGGGACAGAGGAAGCAATATGGTGGGGATTTCCTGAGGGCC AGGATGTCTCTCCAGCACTGACGGCAGGTGCTTCAGGAAAGGTGATGGACTTCAACAAT GGCACCTACCTGGTCAGCTTCACTCTGTTCTGGGAGGGCCAGGTCTCCCTGTCTCTGCTG CTCATCCACCCAGTGAAGGGCGTCGGCTCTCTGGAGGGCAAGGAACCAAGGCTATGAT AAAATTATTTTCAAAGGCAAATTTGTTAATGGCACCTCTCATGTCTTCACTGAATGTGGC CTGACCCTAACTCAAATGCTGAACCTCTGTGAATATCTGGATGACAGAGACCAAGAAGCC TTCTATTGTATGAAGCCTCAACACATGCCCTGTGAGGCTCTGACCTACATGACCACCCGG AATAGAGAGGTATCTTATCTTACAGACAAGGAAAACAGCCTTTCCACAGGTCCAAAGTG GGAGTTGAAATGATGAAGGATCGTAAACACATTGATGTCTACTAATTGTAACAAGAGAGAA AAAATAGAAGAGACATGCCAAGTTGGAATGAAGCCTCCTGTCCCTGGTGGTTATACTTTA CAAGGAAAATGGATAACAACATTTTGCAACCAGGTTCACTTAGACACAATTAAGATAAAT GGCTGTTTGAAAGGCAAACCTCATTTACCTCTGGGAGACTCTACACTACGTCAGTGGATC TACTACTTCCCCAAAGTTGTAAAAACACTGAAGTTTTTGGATCTTCACTGAACATGGAAATC TTTAAGAAACATTTGCTTCTGGATGCAGAAAGACACACTCAGATTCAATGGAAAAACAT AGCTATCCCTTCGTCACTTCCAGCTTACTCTCTGATAGATCATGATTATATCCCTCGG GAAATTGACCGGTATCAGGTGACAAAAACACAGCCATCGTCATCACCTTTGGCCAGCAC TTTAGACCATTTCCCATTGACATTTTATTTCGAGGGCCATCGGTGTTCAAAGGCTATT GAAAGACTGTTCTAAGAAGCCAGCCACTAAAGTGATTATTAAGACAGAAAAACATCAGG GAGATGCACATAGAGACAGAGAGGTTTGGAGACTTCCATGGTTATATTCACTATCTTATC ATGAAGGATATTTCAAAGACCTCAACGTGGGCATCATTGATGCCTGGGACATGACCATT GCATATGGCACTGACACTATCCACCACCTGATCATGTGATTGGAAATCAGATTAAACATG		

	TTCTTAAACTACATTTGCTAAGGGATAAATACTATACAAAATCACTAGGAACCAATCTCT GCACATAATCCACATGTATTGTAAAGTAAGTTTACTCATTTTAGGAACCTAAGGAAAAT AAATTTAAAAGAATCTGTTTGGGGAGGAAGGCTATGTAAGGACAATGACAACCTGATAAGG GATGCAAAACCAAGAGAATCATTCATGAAGAATGACTATACCATGCCTGGTTCTGATGCT CGTTTAAAATATTAAAAAAGTTTTT		
	ORF Start: ATG at 13		ORF Stop: TAA at 1639
	SEQ ID NO: 50	542 aa	MW at 62656.8kD
NOV15a, CG138606-01 Protein Sequence	MLQKTLILISFSVVTWMIFIISQNFKLWSALNLSISVHYWNSAKSLFPKTSLIPLKP LTETELRIKEIEKLDQIPRPPTHVNTTTSATHSTATILNPRDTCRGDQLDILLEVR DHLGQRKQYGGDFLRARMSSPALTAGASGKVMDFNNGTYLVSTLFWEGQVSLSLLIHP SEGASALWRARNQGYDKIIFKGKFNVTSHVFTECGLTLNSNAELCEYLDDRDQEAFCM KPQHMPCEALTYMTTRNREVSYLTDKENS LFHRSKVGVEMMKDRKHIDVTNCNKREKIEE TCQVGMKPPVPGGYTLQGWITTFQCNQVQLDTIKINGCLKGKLIYLLGDSTLRQWIYYFP KVVKTLKFFDLHETGIFKKHLLD AERHTQIQWKKHSYPVTFQLYSLIDHDYIPREIDR LSGDKNTAIVITFGQHFRRPFIIDIFIRRAIGVQKAIERLFLRSPATKVIKKTENIREMHI ETERFGDFHGYIHYLIMKDIKDLNVGIIDAWDMTIAYGTDTHPPDHVIGNQINMFLNY IC		

Further analysis of the NOV15a protein yielded the following properties shown in Table 15B.

Table 15B. Protein Sequence Properties NOV15a	
PSort analysis:	0.6850 probability located in plasma membrane; 0.6400 probability located in endoplasmic reticulum (membrane); 0.3700 probability located in Golgi body; 0.2923 probability located in microbody (peroxisome)
SignalP analysis:	Cleavage site between residues 19 and 20

- A search of the NOV15a protein against the Geneseq database, a proprietary
- 5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 15C.

Table 15C. Geneseq Results for NOV15a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV15a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU96185	Human secreted protein, SEQ ID No 87 - <i>Homo sapiens</i> , 547 aa. [WO200224721-A1, 28-MAR-2002]	1..542 6..547	542/542 (100%) 542/542 (100%)	0.0
ABG27904	Novel human diagnostic protein #27895 - <i>Homo sapiens</i> , 590 aa. [WO200175067-A2, 11-OCT-2001]	26..542 74..590	515/517 (99%) 515/517 (99%)	0.0

AAU83597	Human PRO protein, Seq ID No 12 - <i>Homo sapiens</i> , 544 aa. [WO200208288-A2, 31-JAN-2002]	4..542 9..544	372/540 (68%) 441/540 (80%)	0.0
AAU96219	Human secreted protein, SEQ ID No 121 - <i>Homo sapiens</i> , 303 aa. [WO200224721-A1, 28-MAR-2002]	1..298 6..303	291/298 (97%) 291/298 (97%)	e-170
AAB74709	Human membrane associated protein MEMAP-15 - <i>Homo sapiens</i> , 277 aa. [WO200112662-A2, 22-FEB-2001]	4..273 9..277	220/270 (81%) 245/270 (90%)	e-129

In a BLAST search of public sequence databases, the NOV15a protein was found to have homology to the proteins shown in the BLASTP data in Table 15D.

Table 15D. Public BLASTP Results for NOV15a				
Protein Accession Number	Protein/Organism/Length	NOV15a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q05004	Brush border 61.9 kDa protein precursor - <i>Oryctolagus cuniculus</i> (Rabbit), 540 aa.	1..542 1..540	427/542 (78%) 486/542 (88%)	0.0
AAH29049	Hypothetical 46.9 kDa protein - <i>Homo sapiens</i> (Human), 405 aa.	138..542 1..405	404/405 (99%) 404/405 (99%)	0.0
Q9CX72	4432416J03Rik protein - <i>Mus musculus</i> (Mouse), 558 aa.	6..542 24..558	339/539 (62%) 416/539 (76%)	0.0
Q96DL1	CDNA FLJ25224 fis, clone STM00905 - <i>Homo sapiens</i> (Human), 365 aa.	2..292 18..308	205/291 (70%) 239/291 (81%)	e-116
Q969Y0	CDNA FLJ30102 fis, clone BNGH41000137, weakly similar to brush border 61.9 kDa protein precursor (Unknown) (Protein for MGC:15606) - <i>Homo sapiens</i> (Human), 559 aa.	18..542 19..555	168/543 (30%) 287/543 (51%)	3e-69

PFam analysis predicts that the NOV15a protein contains the domains shown in

5 Table 15E.

Table 15E. Domain Analysis of NOV15a			
Pfam Domain	NOV15a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 16.

The NOV16 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 16A.

Table 16A. NOV16 Sequence Analysis			
	SEQ ID NO: 51	1638 bp	
NOV16a, CG138751-01 DNA Sequence	ACACGCGCCAGCTCTGTAGCCTCCTCCGTCGACTCAGCCTTAGGTACCGGTACAGGCAAA ATGCGGTCTCTCCCTGGCTCCGGAGTCTGGTTCTTCCGGGCCTTCTCCAGGGACAGCTGG TTCCGAGGCCTCATCTGCTGCTGACCTTCCTAATTTACGCCTGCTATCACATGTCCAGG AAGCCTATCAGTATCGTCAAGAGCCGTCTGCACCAGAACTGCTCGGAGCAGATCAAACCC ATCAATGATACTCACAGTCTCAATGACACCATGTGGTGCAGCTGGGCCCCATTTGACAAG GACAACTATAAGGAGTTACTAGGGGGCGTGGACAACGCCTTCCTCATCGCCTATGCCATC GGCATGTTTCATCAGTGGGGTTTTTGGGGAGCGGCTTCCGCTCCGTTACTACCTCTCAGCT GGAATGCTGCTCAGTGGCCTTTTACCTCGCTCTTTGGCCTGGGATATTCTGGAACATC CACGAGCTCTGGTACTTTGTGGTCATCCAGGTCTGTAATGGAATCGTCCAGACCACAGGC TGGCCCTCTGTGGTGACCTGTGTTGGCAACTGGTTCGGGAAGGGGAAGCGGGGTTTCATC ATGGGCATCTGGAATCCCAACACATCTGTGGGCAACATCTGGGCTCCCTGATCGCCGGC ATCTGGGTGAACGGGCAGTGGGGCCTGTCTGTTTCATCGTGCCTGGCATCATTACTGCCGTC ATGGGCGTCATCACCTTCTCTCTCATCGAACACCCAGAAGATGTGGACTGCGCCCT CCTCAGCACCGGTGAGCCAGCTGAGAACCAGGACAACCCTGAGGACCTTGGGAACAGT CCCTGCTCTATCAGGGAGAGCGGCTTGAGACTGTGGCCAAATGCTCCAAGGGGCCATGC GAAGAGCCTGCTGCCATCAGCTTCTTTGGGGCGCTCCGGATCCCAGGCGTGGTTCGAGTTC TCTCTGTGTCTGCTGTTTGCCAAGCTGGTCAGTTACACCTTCTCTACTGGCTGCCCTC TACATCGCCAATGTGGCTCACTTTAGTGCCAAGGAGGCTGGGGACCTGTCTACACTCTTC GATGTTGGTGGCATCATAGGCGGCATCGTGGCAGGGCTCGTCTCTGACTACACCAATGGC AGGGCCACCCTTGCTGTGTATGCTCATCTTGGCTGCCCCCATGATGTTCTGTACAAC TACATTGGCCAGGACGGGATTGCCAGCTCCATAGGTGAGGTCCCAGTGATGCTGATCATC TGTGGGGGCTGGTCAATGGCCCATACGCGCTCATCACCCTGCTGTCTCTGCTGATCTG GGGACTCACAAGAGCCTGAAGGGCACAGCCAAAGCCCTGTCCACGGTCACGGCCATCATT GACGGCACCGGCTCCATAGGTGCGGCTCTGGGGCCTCTGCTGGCTGGGCTCATCTCCCC ACGGGCTGGAACAATGTCTTCTACATGCTCATCTCTGCCGACGTCTAGCCTGCTGGTTC CTTTGCCGGTTAGTATACAAAGAGATCTTGGCCTGGAAGGTGTCCCTGAGCAGAGGCAGC GGGTGAGTCCGGGGAGCTGAAGCTGCCCCCTTACCAACCTCATTCTCGTGGGAATCAGC CCAGCGCTCAGTTTCTCC		
	ORF Start: ATG at 61		ORF Stop: TGA at 1564
	SEQ ID NO: 52	501 aa	MW at 54257.6kD
NOV16a, CG138751-01 Protein Sequence	MRSSLAPGVWFFRAFSRDSWFRGLILLTFLIYACYHMSRKPISIVKSRLHQCSEQIKP INDTHSLNDTMWCSWAPFDKDYKELLGGVDNAFLIAYAIGMFISGVGERLPLRYLSA GMLLSGLFTSLFGLGYFVNIHELWYFVVIQVCNGLVQTTGWPSVVTGVNWFVGKGRGFI MGIWNSHTSVGNILGSLIAGIWNQWGLSFIVPGIITAVMGVITFLFLIEHPEDVDCAP POHHGEPAENQDNPDPGNSPCSIRESGLETVAKCSKGPCEEPAAISFFGALRIPGVVEF SLCLLFAKLVSYTFLYWLPYIANVAHFSAKEAGDLSTLFDVGGIIGGIVAGLVSDYTNL RATTCVMLILAAPMMFLYNIQDGIASSIGVEPVMIIICGLVNGPYALITTAVSADL GTHKSLKGTAKALSTVTAIIDGTGSIGAAALGPLLAGLISPTGWNVNFYMLISADVLACL LCRLVYKEILAWKVSLSRGSG		

	SEQ ID NO: 53	1573 bp	
NOV16b, CG138751-02 DNA Sequence	GACTCAGCCTTAGGTACCGGTCAGGCAAAATGCGGTCTCCCTGGCTCCGGGAGTCTGGT TCTTCCGGGCCTTCTCCAGGGACAGCTGGTTCCGAGGCCTCATCCTGCTGCTGACCTTCC TAATTTACGCCTGCTATCACATGTCCAGGAAGCCTATCAGTATCGTCAAGAGCCGTCTGC ACCAGAACTGCTCGGAGCAGATCAAACCCATCAATGATACTCACAGTCTCAATGACACCA TGTGGTGACAGCTGGGCCCCATTGACAAGGACAACATAAAGAGTTACTAGGGGGCGTGG ACAACGCCTTCTCATCGCCTATGCCATCGGCATGTTTCATCAGTGGGGTTTTGGGGAGC GGCTTCCGCTCCGTTACTACCTCTCAGCTGGAATGCTGCTCAGTGGCCTTTTACCTCGC TCTTTGGCCTGGGATATTTCTGGAACATCCACGAGCTCTGGTACTTTGTGGTCATCCAGG TCTGTAATGGACTCGTCCAGACCACAGGCTGGCCCTCTGTGGTGACCTGTGTTGGCAACT GGTTCGGAAGGGGAAGCGGGGGTTCATCATGGGCATCTGGAATTCCCACACATCTGTGG GCAACATCCTGGGCTCCCTGATCGCCGGCATCTGGGTGAACGGGCAGTGGGCCTGTGCT TCATCGTGCCTGGCATCATTACTGCCGTATGGGCGTCATGAGCCTTCTCTTCTCATCG AACACCCAGAAGATGTGGACTGCGCCCTCCTCAGCACCACGGTGAGCCAGCTGAGAACC AGGACAACCCTGAGGACCCTGGGAACAGTCCCTGCTCTATCAGGAGAGCGGCCTTGAGA CTGTGGCCAAATGTCCAAGGGGCGATGCGAAGAGCCTGCTGCCATCAGCTTCTTTGGGG CGCTCCGGATCCCAGGCGTGGTCTGAGTTCTCTGTGTCTGCTGTTGCCAAGCTGGTCA GTTACACCTTCTCTACTGGCTGCCCTCTACATCGCCAATGTGGCTCACTTTAGTGCCA AGGAGGCTGGGGACCTGTCTACACTCTTCGATGTTGGTGGCATCATAGGCGGCATCGTGG CAGGGCTCGTCTCTGACTACCAATGGCAGGGCCACCATTGCTGTGTATGCTCATCT TGGCTGCCCCCATGATGTTCTGTACAACTACATTGGCCAGGACGGGATTGCCAGCTCCA TAGTGATGCTGATCATCTGTGGGGGCTGGTCAATGGCCCATACGCGCTCATCACCCTG CTGTCTCTGCTGATCTGGGGACTCACAGAGCCTGAAGGGCAACGCCAAGCCCTGTCCA CGGTACGGCCATCATTGACGGCACCGGCTCCATAGGTGCGGCTCTGGGCGCTCTGCTGG CTGGGCTCATCTCCCCACGGGCTGGAACAATGTCTCTACATGCTCATCTCTGCCGACG TCCTAGCCTGCTTGCTCCTTTGCCGGTTAGTATACAAAGAGATCTTGGCCTGGAAGGTGT CCCTGAGCAGAGGAGCGGGTGAAGTCCGGGGAGCTGAAGCTGCCCCCTTACCAACCTCAT TTCTCGTGGGAAT		
	ORF Start: ATG at 30		ORF Stop: TGA at 1521
	SEQ ID NO: 54	497 aa	MW at 53902.2kD
NOV16b, CG138751-02 Protein Sequence	MRSSLAPGVWFFRAFSRDSWFRGLILLTFLIYACYHMSRKPISIVKSLRHQNCSEIQKP INDTHSLNDTMWCSWAPFDKDYKELGGVDNAFLIAYAIGMFISGVFGERLPLRYLISA GMLLSGLFTSLFGLGYFWNIHELWYFVVIQVCNGLVQTTGWPSVVTVCVGNWFGKGRGFI MGIWNSHTSVGNILGSLIAGIWNQWGLSFIVPGIITAVMGVITFLFLIEHPEDVDCAP PQHHGEPAENQDNPEDPGNSPCSIRESGLETVAKCSKGPCEEPAAISFFGALRIPGVVEF SLCLLFAKLVSYTFYWLPLYIANVAHPSAKEAGDLSTLFDVGGIIGGIVAGLVSDYTNG RATTCVMLILAAPMMFLYNYIGQDGIASSIVMLIICGGLVNGPYALITTAVSADLGTHK SLKGNALSTVTAIIDGTGSIGAALGPLLAGLISPTGWNNVFYMLISADVLACLLCRL VYKEILAWKVSLSRGSG		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 16B.

Table 16B. Comparison of NOV16a against NOV16b.		
Protein Sequence	NOV16a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV16b	1..501 1..497	450/501 (89%) 451/501 (89%)

Two polymorphic variants of NOV16a have been identified and are shown in Table 41D. Further analysis of the NOV16a protein yielded the following properties shown in Table 16C.

Table 16C. Protein Sequence Properties NOV16a	
PSort analysis:	0.6318 probability located in mitochondrial inner membrane; 0.6000 probability located in plasma membrane; 0.4778 probability located in mitochondrial intermembrane space; 0.4262 probability located in mitochondrial matrix space
SignalP analysis:	Cleavage site between residues 37 and 38

A search of the NOV16a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 16D.

Table 16D. Geneseq Results for NOV16a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV16a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM00776	Human bone marrow protein, SEQ ID NO: 139 - <i>Homo sapiens</i> , 211 aa. [WO200153453-A2, 26-JUL-2001]	181..391 1..211	205/211 (97%) 206/211 (97%)	e-118
AAM00889	Human bone marrow protein, SEQ ID NO: 365 - <i>Homo sapiens</i> , 201 aa. [WO200153453-A2, 26-JUL-2001]	170..368 3..201	193/199 (96%) 195/199 (97%)	e-113
AAG31980	<i>Arabidopsis thaliana</i> protein fragment SEQ ID NO: 38498 - <i>Arabidopsis thaliana</i> , 476 aa. [EP1033405-A2, 06-SEP-2000]	24..489 31..462	220/470 (46%) 296/470 (62%)	e-110
AAB42327	Human ORFX ORF2091 polypeptide sequence SEQ ID NO:4182 - <i>Homo sapiens</i> , 192 aa. [WO200058473-A2, 05-OCT-2000]	295..489 2..192	185/195 (94%) 187/195 (95%)	e-100
ABB64855	<i>Drosophila melanogaster</i> polypeptide SEQ ID NO 21357 - <i>Drosophila melanogaster</i> , 432 aa. [WO200171042-A2, 27-SEP-2001]	145..491 80..421	192/352 (54%) 232/352 (65%)	4e-98

In a BLAST search of public sequence databases, the NOV16a protein was found to have homology to the proteins shown in the BLASTP data in Table 16E.

Table 16E. Public BLASTP Results for NOV16a				
Protein Accession Number	Protein/Organism/Length	NOV16a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q8TED4	CDNA FLJ23627 fis, clone ADSU02391, highly similar to <i>Mus musculus</i> cAMP inducible 2 protein (Ci2) mRNA - <i>Homo sapiens</i> (Human), 501 aa.	1..501 1..497	494/501 (98%) 495/501 (98%)	0.0
Q9WU81	cAMP inducible 2 protein - <i>Mus musculus</i> (Mouse), 501 aa.	1..501 1..497	435/501 (86%) 461/501 (91%)	0.0
Q8TEM2	FLJ00171 protein - <i>Homo sapiens</i> (Human), 396 aa (fragment).	1..346 12..357	346/346 (100%) 346/346 (100%)	0.0
Q8R070	Similar to solute carrier family 37 (glycerol-3-phosphate transporter), member 1 - <i>Mus musculus</i> (Mouse), 531 aa.	5..489 4..515	308/516 (59%) 377/516 (72%)	e-173
AAF46705	CG10069-PA - <i>Drosophila melanogaster</i> (Fruit fly), 516 aa.	17..491 30..505	257/489 (52%) 320/489 (64%)	e-136

PFam analysis predicts that the NOV16a protein contains the domains shown in
5 Table 16F.

Table 16F. Domain Analysis of NOV16a			
Pfam Domain	NOV16a Match Region	Identities/ Similarities for the Matched Region	Expect Value
sugar_tr	9..494	66/553 (12%) 308/553 (56%)	0.28

Example 17.

The NOV17 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 17A.

Table 17A. NOV17 Sequence Analysis			
	SEQ ID NO: 55	5590 bp	
NOV17a, CG139062-01 DNA Sequence	CTGCGGCCGCGCCGCGAGCTAGGCTGGGTTTTTTTTTTCTCCCTCCCTCCCCCTTTT		
	TCCATGCAGCTGATCTAAAAGGGAATAAAAGGCTGCGCATAATCATAATAATAAAAGAAAG		
	GGGAGCGCGAGAGAAGGAAAGAAAGCCGGGAGGTGGAAGAGGAGGGGAGCGTCTCAAAG		
	AAGCGATCAGAATAATAAAAGGAGGCCGGGCTCTTTCCTTCTGGAACGGGCGGCTCTTG		
	AAAGGGCTTTTGAAAAGTGGTGTGTTTTCCAGTCGTGCATGCTCCAATCGGCGGAGTAT		
	ATTAGAGCCGGGACGCGGCGGCCGAGGGGACGCGGCGACGGCAGCACCGGCGGCGAGCAC		
	CAGCGCGAACAGCAGCGGCGGCGTCCCGAGTGCCCGCGGCGCGGCGGCGAGCGATGCGTT		
	CCCCACGACGCGCGGCGGTCGGGCGCCCCCTAAGCCTCCTGCTCGCCCTGCTCTGTG		
	CCCTGCGAGCCAAGGTGTGTGGGCGCTCGGGTCAGTTCGAGTTGGAGATCCTGTCCATGC		
	AGAACGTGAACGGGAGCTGCAGAACGGGAAGTGTGCGGCGGCGCCCCGAACCCGGGAG		
	ACCGCAAGTGCACCCGCGACGAGTGTGACACATACTTCAAAGTGTGCCTCAAGGAGTATC		
	AGTCCCGCGTCACGCGCGGGGGCCCTGCAGCTTCGGCTCAGGGTCCACGCTGTCTATCG		
	GGGGCAACACCTTCAACCTCAAGGCCAGCGCGGCAACGACCGCAACCGCATCGTGTGCTG		
	CTTTCAGTTTCGCTGGCCGAGGTCTTATACGTTGCTTGTGGAGGCGTGGGATTCCAGTA		
	ATGACACCGTTCAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGCGATGATCAACC		
	CCAGCCGGCAGTGGCAGACGCTGAAGCAGAACACGGGCGTTGCCACTTTGAGTATCAGA		
	TCCGCGTGACCTGTGATGACTACTATGGCTTTGGCTGCAATAAGTTCTGCCGCCCA		
	GAGATGACTTCTTTGGACACTATGCCTGTGACCAGAATGGCAACAAACTTGCATGGAAG		
	GCTGGATGGGCCCCGAATGTAAACAGAGCTATTTGCCGACAAGGCTGCAGTCTTAAGCATG		
	GGTCTTGCAAACCTCCAGGTGACTGCAGGTGCCAGTATGGCTGGCAAGGCGCTGACTGTG		
	ATAAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAATGAGCCCTGGCAGTGCC		
	TCTGTGAGACCAACTGGGCGGCCAGCTCTGTGACAAAGATCTCAATTACTGTGGGACTC		
	ATCAGCCGTGTCTCAACGGGGGAAGTGTAGCAACACAGGCCCTGCAAAATATCAGTGT		
	CCTGCCCTGAGGGGTATTTCAGGACCCAAGTGTGAAATTGCTGAGCACGCTGCCTCTCTG		
	ATCCCTGTCAACAGAGGCGCTGTAAGGAGACCTCCCTGGGCTTTGAGTGTGAGTGT		
	CCCCAGGCTGGACCGGCCCCACATGCTCTACAAACATTGATGACTGTTCTCCTAATAACT		
	GTTCCACGGGGGACCTGCCAGGACCTGGTTAACGGATTAAAGTGTGTGTGCCCCCAC		
	AGTGGACTGGGAAAACGTGCCAGTTAGATGCAAAATGAATGTGAGGCCAAACCTTGTGTAA		
	ACGCCAAATCCTGTAAGAATCTCATTGCCAGCTACTACTGCGACTGTCTCCGGCTGGA		
	TGGGTCAGAATTGTGACATAAATATTAATGACTGCCTTGGCCAGTGTGAGAATGACGCT		
	CCTGTGCGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGGCTATGCAGGCGATC		
	ACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTGAATGGGGTCACTGTG		
	AGAAATGAAATCAACAGATTCCAGTGTCTGTGTCCACTGGTTCTCTGGAAACCTGTGTC		
	AGCTGGACATCGATTATTGTGAGCCTAATCCCTGCCAGAACGGTGCCAGTGCTACAACC		
	GTGCCAGTGACTATTTCTGCAAGTGCCCGAGGACTATGAGGGCAAGAACTGCTCACACC		
	TGAAAGACCACTGCCGACGACCCCTGTGAAGTGATTGACAGCTGCACAGTGGCCATGG		
	CTTCCAACGACACACCTGAAGGGGTGCGGTATATTTCTCCAACGTCTGTGGTCTCACG		
	GGAAGTGCAAGAGTCAGTCGGGAGGCAAAATTCACCTGTGACTGTAACAAAGCTTCACGG		
	GAACATACTGCCATGAAAATATTAATGACTGTGAGAGCAACCCCTGTAGAAACGGTGGCA		
	CTTGATCGATGGTGTCAACTCCTACAAGTGCATCTGTAGTGACGGCTGGGAGGGGCGCT		
	ACTGTGAAACCAATATTAATGACTGCAGCCAGAACCCCTGCCACAATGGGGGACGTGTC		
	GCGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGGTGGAAAGGAAAGACCTGCC		
	ACTACGTGACAGTCAGTGTGATGAGGCCACGTGCAACAACGGTGGCACCTGCTATGATG		
	AGGGGGATGCTTTTAAGTGCATGTGTCTGGCGGCTGGGAAGGAACAACCTGTAACATAG		
	CCCGAAACAGTAGCTGCCTGCCCAACCCCTGCCATAATGGGGGCACATGTGTGGTCAACG		
	GCGAGTCCTTTACGTGCGTCTGCAAGGAAGGCTGGGAGGGGCCATCTGTGCTCAGAATA		
	CCAATGACTGCAGCCCTCATCCCTGTTACAACAGCGGCACCTGTGTGGATGGAGACAACT		
	GGTACCGGTGCGAATGTGCCCCGGGTTTTGCTGGGCCCCACTGCAGAATAAACATCAATG		
	AATGCCAGTCTTCACCTTGTGCCTTTGGAGCGACCTGTGTGGATGAGATCAATGGCTACC		
	GGTGTGTCTGCCCTCCAGGGCACAGTGGTGCCAAGTGCCAGGAAGTTTCAGGGAGACCTT		
	GCATCACCATGGGGAGTGTGATACCAGATGGGGCCAAATGGGATGATGACTGTAATACCT		
	GCCAGTGCCTGAATGGACGGATCGCTGCTCAAAGGTCTGGTGTGGCCCTCGACCTTGCC		
	TGCTCCACAAAGGGCACAGCGAGTGCCCCAGCGGGCAGAGCTGCATCCCCATCTTGACG		
	ACCAAGTCTTGTCCACCCCTGCAGTGGTGTGGGCGAGTGTGCGTCTTCCAGTCTCCAGC		
	CGGTGAAGACAAAGTGCACCTCTGACTCCTATTACCAGGATAACTGTGCGAACATCACAT		
	TTACCTTTAACAAGGAGATGATGTACCAGGTCTTACTACGAGCACATTTGCAGTGAAT		
	TGAGGAATTTGAATATTTTGAAGAATGTTTCCGCTGAATATTCAATCTACATCGCTTGGC		

	AGCCTTCCCTTCAGCGAACAATGAAATACATGTGGCCATTTCTGCTGAAGATATACGGG ATGATGGGAACCCGATCAAGGAAATCACTGACAAAATAATCGATCTTGTTAGTAAACGTG ATGGAAACAGCTCGCTGATTGCTGCCGTTGCAGAAAGTAAGAGTTCAGAGGCGGCCTCTGA AGAACAGAACAGATTTCTTGTTCCTTGTCTGAGCTCTGTCTTAACGTGGCTTGGATCT GTTGCTTGGTGACGGCCTTCTACTGGTGCCTGCGGAAGCGGCGGAAGCGGACCCACA CACACTCAGCCTCTGAGGACAAACACCACCAACAACGTGCGGGAGCAGCTGAACCAGATCA AAAACCCCATTTAGAAACATGGGGCCAACACGGTCCCATCAAGGATTACGAGAACAAGA ACTCCAAATGTCTAAATAAGGACACACAATTCTGAAGTAGAAGAGGACGACATGGACA AACACCAGCAGAAAGCCCGTTTGCCAAGCAGCCGCGTATACGCTGGTAGACAGAGAAG AGAAGCCCCCAACGGCACGCCGACAAAACCCAACTGGACAAAACAGGACAACA GAGACTTGAAAGTGCCAGAGCTTAAACCGAATGGAGTACATCGTATATAGCAGACCGCG GCACTGCCGCGCTAGGTAGAGTCTGAGGGCTTGTAGTTCTTTAACTGTCTGTCTATAC TCGAGTCTGAGGCGTTGCTGACTTAGAATCCCTGTGTTAATTTAAGTTTTGACAAGCTG GCTTACACTGGCAATGGTAGTTTCTGTGGTTGGCTGGGAAATCGAGTGCCGCATCTCACA GCTATGCAAAAAGCTAGTCAACAGTACCCTGGTTGTGTGTCCTTGCAGCCGACACGGT CTCGGATCAGGCTCCAGGAGCCTGCCAGCCCCCTGGTCTTTGAGCTCCCACTTCTGCC AGATGTCCTAATGGTGATGCAGTCTTAGATCATAGTTTTATTTATATTTTACTCTTG AGTTGTTTTGTATATTGGTTTTATGATGACGTACAAGTAGTTCTGTATTGAAAGTGCC TTTGAGCTCAGAACACAGCAACGATCACAATGACTTTATTTATTTATTTTAAATG TATTTTTGTGTTGGGGAGGGGAGACTTTGATGTCAGCAGTTGCTGGTAAATGAAGAA TTTAAAGAAAAAATGTCAAAGTAGAACTTTGTATAGTTATGTAAATAATCTTTTTTA TTAATCACTGTGTATATTTGATTATTAACCTAATAATCAAGAGCCTTAAACATCATTC CTTTTTATTTATATGTATGTGTTTAGAATGAAGGTTTTTGATAGCATTGTAAGCGTATG GCTTTATTTTTTGAACCTCTCTCATTACTTGTGCTATAAGCAAAATTAAGGTGTTT GAAAATAGTTTATTTTAAACAATAGGATGGGCTTCTGTGCCAGAATCTGATGGAATT TTTTTTGTACGACGTGAGATGTTTAAACACCTTCTATAGCATCACTTAAACACGTTTT AAGGACTGACTGAGGCAGTTTGAAGATTAGTTTAGAACAGGTTTTTTTGTGTTGTTT TTTGTGTTTCTGCTTTAGACTTGAAAGAGACAGGCAGGTGATCTGCTGCAGAGCAGTAA GGGAACAAGTTGAGCTATGACTTAACATAGCCAAATGTGAGTGGTTGAATATGATTAAA AATATCAAATTAATTGTGTGAACCTGGAAGCACACCAATCTGACTTTGTAATTTCTGATT TCTTTTACCATTGCTACATAACTGAACCACTTGTAGATTGATTTTTTTTAAATCT ACTGCATTAGGGAGTATTCTAATAAGCTAGTTGAATACTTGAACCATAAAATGTCCAGT AAGATCACTGTTTAGATTTGCCATAGAGTACACTGCCTGCCTTAAGTGAGGAAATCAAAG TGCTATTACGAAGTTCAAGATCAAAAAGGCTTATAAAACAGAGTAATCTTGTGTTTAC CATTGAGACCGTGAAGATACTTTGTATTGTCTATTAGTGTTATATGAACATACAAATGC ATCTTTGATGTGTTGTTCTTGCAATAAATTTGAAAAGTAATTTTATTAATTTTTTT GTATGAAAAC		
	ORF Start: ATG at 414		ORF Stop: TAG at 4068
	SEQ ID NO: 56	1218 aa	MW at 133797.1kD
NOV17a, CG139062-01 Protein Sequence	MRSPTTRGRSGRPLSLLLALLCALRAKVCASGQFELEILSMQNVNGELQNGNCCGGARN PGDRKCTRDECPTYFKVCLKEYQSRVTAGGPCSFSGSTPVIIGNTFNLKASRGNDNRRI VLPFSFAWPRSYLLVEAWDSSNDTVQPDSEIEKASHSGMINPSRQWTLKQNTGVAHFE YQIRVTCDDYYYGFGCNKFCRPRDDFFGHYACDQNGNKTCEGMWGMPECNRAICRQGCSP KHGSKLPDGRCCQYWGQGLYCDKCIHPHGCVHIGICNEPWQCLCETNWWGQLCDKDLNYC GTHQPCLNNGGTCSTNTPDKYQCSCEPGYSNPCEIAEHACLSDPCHNRGSCKETSGLFEC ECSPGWTGPTCSTNIDDCSPNNCSHGTCQDLVNGFKVCPPQWTGKTCQLDANECEAKP CVNAKSKNLIASYYCDCLPGWMQNCNDININDCLGQCQNDASCRDLVNGYRICPPGYA GDHCERDIDECASNPLNGGHCQNEINRFQCLCPTGFSNLCQLDIDYCEPNPCQNGAQC YNRASDYFCKCPEDYEGKNCSHLKDHCRTPCEVIDSCTVAMASNDTPGEVRYISSNVCG PHGKCKSQSGGKFTCDCKNGFTGTCHENINDCESNPCRNNGTCTIDGVNSYKICSDGWE GAYCETNINDCSQNPCHNGGTCRDLVNDYCDCKNGWKGTCHSRDSQCDEATCNGGTC YDEGDAFKCMCPGGWEGTTCNIARNSSCLPNPCHNGGTCVVNGESFTVCCKEGWEGPICA QNTNDCSPHPCYNSGTCDVDGNWYRCECAPGFAGPDCRININECQSSPCAFCATCDEIN GYRCVCPPGHSGAKQEVSGRPCITMGSVIPDGAkWDDDCNTCQCLNLRGACSKVWCGR PCLLHKHSECPSPGQSCIPILDDQCFVHPCTGVGECRSSSLQPVKTKCTSDSYQDNCAN ITFTFNKEMMSPGLTTEHICSELNRLNKLNVSAEYSIYIACEFSPSANNEIHVAISAED IRDDGNPIKEITDKIIDLVSKRDGNSSLIAAFAEVRVQRRPLKNRTDFLVPILLSVLTVA WICCLVTAIFYWCLRRKPKGSHTHSASEDNTTNNVREQLNQIKNPIEKHGANTVPIKDYE		

	NKNSKMSKIRTHNSEVEEDMDKHQKARFAKQPAYTLVDREKPPNGTPTKHPNWTNKQ DNRDLESAQSLNRMEYIV		
	SEQ ID NO: 57	4333 bp	
NOV17b, CG139062-02 DNA Sequence	CTGCGGCCGGCCCCGAGCTAGGCTGGGTTTTTTTTTTCTCCCTCCCTCCCCCTTTT TCCATGCAGCTGATCTAAAAGGGAATAAAAGGCTGCGCATAATCATAATAATAAAGAAG GGGAGCGCGAGAGAAGGAAAGAAAGCCGGGAGGTGGAAGAGGAGGGGAGCGTCTCAAAG AAGCGATCAGAATAATAAAAGGAGGCCGGGCTCTTGCCTTCTGGAACGGGCCGCTCTTG AAAGGGCTTTTGAAGTGGTGTGTTTTCCAGTCGTGCATGTCCAATCGGCGGAGTAT ATTAGAGCCGGGACGCGCGGCCGAGGGGAGCGGCGACGGCAGCACCGGCGGCAGCAC CAGCGGAACAGCAGCGCGGCCGTCCCGAGTGCCCGCGGCGCGCGGCGCAGCGATGCGTT CCCCACGGACGCGCGGCCGTCCGGGCGCCCCCTAAGCCTCCTGCTCGCCCTGCTCTGTG CCCTGCGAGCCAAGGTGTGTGGGGCCTCGGGTCAGTTCGAGTTGGAGATCCTGTCCATGC AGAACGTGAACGGGAGCTGCAGAACGGGAAGTGTGCGGCGGCCCGGAACCCGGGAG ACCGCAAGTGCACCCGCGACGAGTGTGACACATACTTCAAAGTGTGCTCAAGGAGTATC AGTCCCGCGTCACGGCCGGGGGGCCCTGCAGCTTCGGCTCAGGGTCCACGCTGTCTATCG GGGGCAACACCTTCAACCTCAAGGCCAGCGCGGCAACGACCGCAACCGCATCGTGTGTG CTTTCAGTTTCGCCGTGGCCGAGGTCTTATACGTTGCTTGTGGAGGCGTGGGATTCCAGTA ATGACACCGTTCAACCTGACAGTATTATTGAAAAGGCTTCTCACTCGGGCATGATCAACC CCAGCCGCGAGTGGCAGACGCTGAAGCAGAACACGGGCGTTGCCCACTTTGAGTATCAGA TCCGCGTGACCTGTGATGACTACTACTATGGCTTTGGCTGCAATAAGTTCTGCCGCCCA GAGATGACTTCTTTGGACACTATGCCTGTGACCAGAATGGCAACAAACTTGCATGGAAG GCTGGATGGGCCCCGAATGTAACAGAGCTATTTGCCGACAAGGCTGCAGTCCCTAAGCATG GGTCTTGCAAACTCCAGGTGACTGCAGGTGCCAGTATGGCTGGCAAGGCTGTACTGTG ATAAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAATGAGCCCTGGCAGTGCC TCTGTGAGACCAACTGGGGCGGCCAGCTCTGTGACAAAGATCTCAATTACTGTGGGACTC ATCAGCCGTGTCTCAACGGGGAACTTTAGCAACACAGGCCCTGACAAATATCAGTGTT CCTGCCCTGAGGGGTATTTCAGGACCCAACTGTGAAATTGCTGAGCACGCTGCCTCTCTG ATCCCTGTCAACAGAGGCAGCTGTAAGGAGACCTCCCTGGGCTTTGAGTGTGAGTGT CCCCAGGCTGGACCGGCCCATGCTCTACAAACATTGATGACTGTTCTCCTAATAACT GTTCCACGGGGGCACCTGCCAGGACCTGGTTAACGGATTTAAGTGTGTGTGCCCCCAC AGTGGACTGGGAAAACGTGCCAGTTAGATGCAATGAATGTGAGGCCAAACCTTGTGTAA ACGCCAAATCCTGTAAGAATCTCATTGCCAGCTACTACTGCGACTGTCTCCCGGCTGGA TGGGTGAGAATTGTGACATAAATATTATGACTGCCTTGGCCAGTGTGAGAATGACGCCT CCTGTGCGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACTGGCTATGAGGCGATC ACTGTGAGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTGAATGGGGTCACTGTC AGAATGAAATCAACAGATTCCAGTGTCTGTGTCCACTGGTTTCTCTGGAACCTCTGTCT AGCTGGACATCGATTATTGTGAGCCTAATCCCTGCCAGAACGGTGCCCACTGTCTACACC GTGCCAGTGACTATTTCTGCAAGTGCCCGGAGGACTATGAGGGCAAGAACTGCTCACACC TGAAAGACCAGTCCCGACGACCCCTGTGAAGTGATTGACAGCTGCACAGTGCCCATGG CTTCCAACGACACACCTGAAGGGGTGCGGTATATTCTCCAACGTCTGTGGTCTCTACG GGAAGTGCAAGAGTCAGTCGGGAGGCAAAATCACCTGTGACTGTAACAAAGGCTTCACGG GAACATACTGCCATGAAAATATTAATGACTGTGAGAGCAACCCCTGTAGAAACGGTGGCA CTTGCAATCGATGGTGTCAACTCTTACAAGTGCATCTGTAGTGACGGCTGGGAGGGGGCCT ACTGTGAAACCAATATTAATGACTGCAGCCAGAACCCCTGCCCAATGGGGGCAGTGTCT GCGACCTGGTCAATGACTTCTACTGTGGCTGTAAAAATGGGTGGAAAGGAAAGACCTGCC ACTCACGTGACAGTCAGTGTGATGAGGCCAACACGGTCCCCATCAAGGATTACGAGAACA AGAACTCCAAAATGTCTAAAATAAGGACACACAATCTGAAGTAGAAGAGGACGACATGG ACAAACACCAGCAGAAAGCCCGGTTTGCCAAGCAGCCGGCGTACACGCTGGTAGACAGAG AAGAGAAGCCCCCAACGGCAGCCGACAAAACACCCAACTGGACAAACAAACAGGACA ACAGAGACTTGAAAGTGCCAGAGCTTAAACCGAATGGAGTACATCGTATAGCAGACCG CGGGCACTGCCGCCGCTAGGTAGAGTCTGAGGGCTGTAGTTCTTTAACTGTCTGTGCA TACTCGAGTCTGAGGCCGTTGCTGACTTAGAATCCCTGTGTTAATTTAAGTTTGGACAAG CTGGCTTACACTGGCAATGGTAGTTTCTGTGGTTGGCTGGGAAATCGAGTGCCGCATCTC ACAGCTATGCAAAAAGCTAGTCAACAGTACCCTGGTGTGTGTGCCCTTGACAGCCGACAC GGTCTCGGATCAGGCTCCAGGAGCTGCCAGCCCTGGTCTTTGAGTCCCACCTCT GCCAGATGTCTAATGGTGATGCAGTCTTAGATCATAGTTTATTTATATTTATTGACTC TTGAGTTGTTTTGTATATTGGTTTTATGATGACGTACAAGTAGTTCTGTATTTGAAAGT GCCTTGCAGCTCAGAACCACAGCAACGATCACAATGACTTTATTATTTATTTTTTAA TTGTATTTTGTGTTGGGGGAGGGGAGACTTTGATGTGACGAGTTGCTGGTAAATGAA		

	GAATTTAAAGAAAAAATGTCAAAAGTAGAACTTTGTATAGTTATGTAAATAATTCTTTT TTATTAATCACTGTGTATATTGATTTTATTAACCTTAATAATCAAGAGCCTTAAACATCA TTCCTTTTATTTATATGTATGTGTTTAGAATTGAAGGTTTTTGATAGCATTGTAAGCGT ATGGCTTTATTTTTTTGAACCTTCTCATTACTTGTGTCCTATAAGCCAAAATTAAGGTG TTTGAAATAGTTTATTTTAAACAATAGGATGGGCTTCTGTGCCAGAACTACTGATGGA ATTTTTTTGTACGACGTCAGATGTTTAAACACCTTCTATAGCATCACTTAAACACGT TTTAAGGACTGACTGAGGCAGTTTGAAGATTAGTTTAGAACAGGTTTTTTTGTGTTGTTG TTTTTTGTTTTCTGCTTTAGACTTGAAAAGAGACAGGCAGGTGATCTGCTGCAGAGCAC TAAGGGAACAAGTTGAGCTATGACTTAACATAGCCAAAATGTGAGTGGTTGAATATGATT AAAAATATCAAATTAATTGTGTGAACCTTGAAGCACACCAATCTGACTTTGTAAATTCTG ATTTCTTTTCACCATTTCGTACATAATACTGAACCACTGTAGATTTGATTTTTTTTTAA TCTACTGCATTTAGGGAGTATTCTAATAAGCTAGTTGAATACTTGAACCATAAAATGTCC AGTAAGATCACTGTTTAGATTGGCCATAGAGTACACTGCCTGCCTTAAGTGAGGAAATCA AAGTGCTATTACGAAGTTCAAGATCAAAAAGGCTTATAAAACAGAGTAATCTTGTGTGTT CACCATTGAGACCGTGAAGATACTTTGTATTGTCTATTAGTGTATATGAACATACAAA TGCATCTTTGATGTGTTGTTCTTGGCAATAAATTTGAAAAGTAATATTATTAATTTT TTTGTATGAAAAAC		
	ORF Start: ATG at 414		ORF Stop: TAG at 2811
	SEQ ID NO: 58	799 aa	MW at 88212.4kD
NOV17b, CG139062-02 Protein Sequence	MRSPRTRGRSGRPLSLLLALLCALRAKVCASGQFELEILSMQNVNGELQNGNCCGGARN PGDRKCTRDECDTYFKVCLKEYQSRVTAGGPCSFSGSTPVIGNTFNLKASRGNDRNR I VLPFSFAWPRSYLLVEAWDSSNDTVQPDSEIEKASHSGMINPSRQWQTLKQNTGVAHFE YQIRVTCDDYYYGFGCNKFCRPRDDFFGHYACDQNGNKTCEGWMGPECNRAICRQCSE KHGSKCLPGDCRCQYQWQGLYCDKCI PHPGCVHVICNEPWQCLCETNWWGQLCDKDLNYC GTHQPCLNNGGTCSTNTPGDKYQCSCPEGYSGPNCIEAETHACLSDFCHNRGSCKETS LGFEC ECS PGWTGPTCSTNIDDCSPNCSHGCTCQDLVNGFKVCVPPQWTGKTCQLDANECEAKE CVNAKSKNLIASYYCDCLPGWMQNCNDININDCLGQCQNDASCRDLVNGYRCICPPGYA GDHCERDIDECASNPLNGGHQNEINRFQCLCPTGFSGNLCQLDIDYCEPNPCQNGAQC YNRASDYFCKCPEDYEGKNCSHLKDHCRTPCEVIDSCTVAMASNDTPEGVRYISSNVCG PHGKCKSQSGGKFTCDCKNGFTGTYPHENINDCESNPCRNGGTCIDGVNSYKICSDGWE GAYCETNINDCSQNPCHNGGTCRDLVNDFYCGCKNGWKGTCHSRDSQCDEANTVPIKDY ENKNSKMSKIRTHNSEVEEDMDKHQQKARFAKQPAYTLVDREKPPNGTPTKHPNWTNK QDNRDLESAQSLNRMEYIV		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 17B.

Table 17B. Comparison of NOV17a against NOV17b.		
Protein Sequence	NOV17a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV17b	27..712	685/686 (99%)
	27..712	685/686 (99%)

Five polymorphic variants of NOV17b have been identified and are shown in Table 41E.

- 5 Further analysis of the NOV17a protein yielded the following properties shown in Table 17C.

Table 17C. Protein Sequence Properties NOV17a

PSort analysis:	0.4600 probability located in plasma membrane; 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in outside
SignalP analysis:	Cleavage site between residues 34 and 35

A search of the NOV17a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 17D.

Table 17D. Geneseq Results for NOV17a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV17a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB07822	Human notch agonist ligand - <i>Homo sapiens</i> , 1218 aa. [WO200218544-A2, 07-MAR-2002]	1..1218 1..1218	1218/1218 (100%) 1218/1218 (100%)	0.0
AAW87894	Human JAGGED1 protein - <i>Homo sapiens</i> , 1218 aa. [WO9858958-A2, 30-DEC-1998]	1..1218 1..1218	1218/1218 (100%) 1218/1218 (100%)	0.0
AAW44301	Human serrate 1 - <i>Homo sapiens</i> , 1218 aa. [WO9802458-A1, 22-JAN-1998]	1..1218 1..1218	1218/1218 (100%) 1218/1218 (100%)	0.0
AAU84344	Protein JAG1 differentially expressed in breast cancer tissue - <i>Homo sapiens</i> , 1218 aa. [WO200210436-A2, 07-FEB-2002]	1..1218 1..1218	1217/1218 (99%) 1217/1218 (99%)	0.0
AAY59597	Human Serrate protein sequence - <i>Homo sapiens</i> , 1218 aa. [US6004924-A, 21-DEC-1999]	1..1218 1..1218	1215/1218 (99%) 1216/1218 (99%)	0.0

5 In a BLAST search of public sequence databases, the NOV17a protein was found to have homology to the proteins shown in the BLASTP data in Table 17E.

Table 17E. Public BLASTP Results for NOV17a
--

Protein Accession Number	Protein/Organism/Length	NOV17a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P78504	Jagged 1 precursor (Jagged1) (hJ1) - <i>Homo sapiens</i> (Human), 1218 aa.	1..1218 1..1218	1218/1218 (100%) 1218/1218 (100%)	0.0
Q9QXX0	Jagged 1 precursor (Jagged1) - <i>Mus musculus</i> (Mouse), 1218 aa.	1..1218 1..1218	1176/1218 (96%) 1194/1218 (97%)	0.0
Q63722	Jagged 1 precursor (Jagged1) - <i>Rattus norvegicus</i> (Rat), 1219 aa.	1..1218 1..1219	1175/1219 (96%) 1191/1219 (97%)	0.0
A56136	jagged protein precursor - rat, 1220 aa.	1..1218 1..1220	1168/1223 (95%) 1184/1223 (96%)	0.0
Q90819	C-Serate-1 protein - <i>Gallus gallus</i> (Chicken), 1193 aa (fragment).	27..1218 1..1193	1047/1193 (87%) 1111/1193 (92%)	0.0

PFam analysis predicts that the NOV17a protein contains the domains shown in Table 17F.

Table 17F. Domain Analysis of NOV17a			
Pfam Domain	NOV17a Match Region	Identities/ Similarities for the Matched Region	Expect Value
DSL	167..229	42/67 (63%) 63/67 (94%)	3.9e-40
EGF	300..333	18/47 (38%) 28/47 (60%)	1e-06
EGF	340..371	16/47 (34%) 26/47 (55%)	3.3e-08
EGF	378..409	18/47 (38%) 30/47 (64%)	2.9e-09
EGF	416..447	13/47 (28%) 19/47 (40%)	0.003
EGF	454..484	14/47 (30%) 26/47 (55%)	4.6e-07
EGF	491..522	16/47 (34%) 24/47 (51%)	1.7e-07
EGF	529..560	17/47 (36%) 26/47 (55%)	2.5e-08

EGF	595..626	13/47 (28%) 24/47 (51%)	0.19
EGF	633..664	15/47 (32%) 25/47 (53%)	1.3e-08
EGF	671..702	15/47 (32%) 30/47 (64%)	1.1e-09
EGF	709..740	13/47 (28%) 23/47 (49%)	0.00072
EGF	748..779	17/47 (36%) 27/47 (57%)	3.1e-09
EGF	786..817	17/47 (36%) 28/47 (60%)	3.5e-07
EGF	824..855	16/47 (34%) 25/47 (53%)	1.7e-05
vwc	863..917	18/84 (21%) 33/84 (39%)	0.055

Example 18.

The NOV18 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 18A.

Table 18A. NOV18 Sequence Analysis			
	SEQ ID NO: 59	587 bp	
NOV18a, CG139363-01 DNA Sequence	GGAGCTTGCTGACCATCCCTGGGAGCTTTAATGTTTACTTCTATCTTGCAGAGTTTTCA CTGAAGTTCACCTGCCGCGAACACAGTAAGTACAGCAGCCCCATTGACATCTGGT AAGGGCGACTGTGGGCCCTCTCTGGATTAGCGGCGGGCATACCATTTGCTGGTGGCCACA GCCCTGCTGGTGGCTTTACTATTTACTTTGATTACCCGAAGAAGAAGCAGCATTGAGGCC ATGGAGGTGATTAGTCCATCTGTATGAAAGAATTCTGCTGTAGTTTTTAAAAAACCT ATTTGTTTCCTTAAGAATCCTAGGAGATCACCCACACATGAGAAGAATACGATGGGAGCA CAAGAGGCCACATATATGTGAAGACTGTAGCAGGAAGCGAGGAACCTGTGCATGACCGT TACCGTCCTACTATAGAAATGGAAGAAGGAGGGGATTGTGGTGGCTTGTGCCAGACTG AGCCTGGAATTGATGCAGCTCAGTCAAGGAGCAGCAGACCTGGCACTGGAACAGGGTTGA AAACCAGGGTTTTGTACTTGGAGAGGAAGATGCCAAGCTGCTTCT		
	ORF Start: ATG at 31		ORF Stop: TGA at 538
	SEQ ID NO: 60	169 aa	MW at 18578.4kD
NOV18a, CG139363-01 Protein Sequence	MFTSILQSFSLNFTLPANTVSTAAPIQTSKGDCGPSLGLAAGIPLLVATALLVALLFTL IHRRSSIEAMEVISPSCKEFSADVFKKPICFKNPRRSPTHEKNTMGAQEAHIYVKTV AGSEEPVHDIRPTIEMERRRGLWLVPRLSLELMQLSQGAADLALAEQG		
	SEQ ID NO: 61	528 bp	
NOV18b, CG139363-02 DNA Sequence	GGGAGCTTTAATGTTTACTTCTATCTTGCAGAGTTTTTCACTGAAGTTCACCTGCCGCG GAACACAACGTCCTCTCCTGTACAGGTGGGAAAGAAACGGACTGTGGGCCCTCTCTTGG ATTAGCGGCGGGCATACCATTTGCTGGTGGCCACAGCCCTGCTGGTGGCTTTACTATTTAC TTTGATTACCCGAAGAAGAAGCAGCATTGAGGCCATGGAGGAAAGTGACAGACCATGTGA AATTCAGAAATTGATGACAATCCCAAGATATCTGAGAATCCTAGGAGATCACCCACACA TGAGAAGAATACGATGGGAGCACAAGAGGCCACATATATGTGAAGACTGTAGCAGGAAG CGAGGAACCTGTGCATGACCGTTACCGTCCTACTATAGAAATGGAAGAAGGAGGGGATT GTGGTGGCTTGTGCCAGACTGAGCCTGGAATGATGCAGCTCAGTCAAGGAGCAGCAGAC		

	CTGGCACTGGAACAGGGTTGAAAACCCAGGGTTTTGTACTTGGAGAGG		
	ORF Start: ATG at 11		ORF Stop: TGA at 452
	SEQ ID NO: 62	147 aa	MW at 16372.4kD
NOV18b, CG139363-02 Protein Sequence	MFTSILQSFSLNFTLPANTTSSPVTGGKETDCGPSLGLAAGIPLL VATALLVALLFTLIH RRRSSIEAMEESDRPCEISEIDNPKISENPRRSPTHEKNTMGAQEAHIYVKT VAGSEEP VH DRYRPTIEMERRRGLWWLVPRLSLE		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 18B.

Table 18B. Comparison of NOV18a against NOV18b.		
Protein Sequence	NOV18a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV18b	1..153	108/153 (70%)
	1..147	114/153 (73%)

Further analysis of the NOV18a protein yielded the following properties shown in Table 18C.

Table 18C. Protein Sequence Properties NOV18a	
PSort analysis:	0.8569 probability located in mitochondrial inner membrane; 0.4456 probability located in mitochondrial intermembrane space; 0.2847 probability located in mitochondrial matrix space; 0.2847 probability located in mitochondrial outer membrane
SignalP analysis:	Cleavage site between residues 64 and 65

- 5 A search of the NOV18a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 18D.

Table 18D. Geneseq Results for NOV18a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV18a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABG23422	Novel human diagnostic protein #23413 - <i>Homo sapiens</i> , 163 aa. [WO200175067-A2, 11-OCT-2001]	8..153 15..163	123/153 (80%) 127/153 (82%)	3e-58
AAM79058	Human protein SEQ ID NO 1720 - <i>Homo sapiens</i> , 141 aa. [WO200157190-A2, 09-AUG-2001]	8..153 2..141	116/146 (79%) 122/146 (83%)	1e-56

AA Y94922	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50 - <i>Homo sapiens</i> , 141 aa. [WO200009552-A1, 24-FEB-2000]	8..153 2..141	115/146 (78%) 121/146 (82%)	1e-55
ABG23423	Novel human diagnostic protein #23414 - <i>Homo sapiens</i> , 209 aa. [WO200175067-A2, 11-OCT-2001]	8..158 35..179	115/151 (76%) 122/151 (80%)	2e-55
AAM80042	Human protein SEQ ID NO 3688 - <i>Homo sapiens</i> , 133 aa. [WO200157190-A2, 09-AUG-2001]	8..141 11..133	104/134 (77%) 109/134 (80%)	3e-47

In a BLAST search of public sequence databases, the NOV18a protein was found to have homology to the proteins shown in the BLASTP data in Table 18E.

Table 18E. Public BLASTP Results for NOV18a				
Protein Accession Number	Protein/Organism/Length	NOV18a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96PE5	Transmembrane protein HTMP10 - <i>Homo sapiens</i> (Human), 141 aa.	8..153 2..141	116/146 (79%) 122/146 (83%)	4e-56
Q29102	Transmembrane protein sp83.5 - <i>Sus scrofa</i> (Pig), 142 aa.	8..153 2..142	104/147 (70%) 117/147 (78%)	5e-50
P54423	Cell wall-associated protease precursor (EC 3.4.21.-) [Contains: Cell wall-associated polypeptides CWBP23 and CWBP52] - <i>Bacillus subtilis</i> , 894 aa.	91..167 662..737	22/77 (28%) 39/77 (50%)	2.7
Q9A7Z7	Hypothetical protein CC1570 - <i>Caulobacter crescentus</i> , 311 aa.	108..151 184..227	14/44 (31%) 23/44 (51%)	3.5
Q8S9L6	AT4g21410/T6K22_140 - <i>Arabidopsis thaliana</i> (Mouse-ear cress), 679 aa.	16..77 265..326	19/62 (30%) 32/62 (50%)	3.5

PFam analysis predicts that the NOV18a protein contains the domains shown in Table 18F.

Table 18F. Domain Analysis of NOV18a			
Pfam Domain	NOV18a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 19.

The NOV19 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 19A.

Table 19A. NOV19 Sequence Analysis			
	SEQ ID NO: 63	471 bp	
NOV19a, CG140188-01 DNA Sequence	CCACCCTTGCTGCCACTAACATGGAGACTTTGTACCGTGTCCCATTCTTAGTGCTCGAAT GTCCCAACCTGAAGCTGAAGAAGCCGCCCTGGCTGCAAGTGCTGTCGGCCATGATTGTGT ATGCTCTGATGGTGGTGTCTTACTTCTCGTCACTGGAGGAATAATTTATGATGTTATTG TTGAACCTCCAAGCATTGGCTCTATGACTGATGAACACGGGCATCAGAGGCCAGTAGCTT TCTTGGCCTACAGAGTAAATGAACAATGTATTATGGAAGGACTTGCATCCAGCTTCCTGT TTACAATAGGAGGTTTAGGTTTCATATTCCTGGACCGATGGAATGCACCAAATATCCCAA AACTCAATAGATTCTTCTTCTATTTCATTGGATTCTGTTGTGTCCTATTGAGCTTTTCA TGGCTAGAGTATTCATGAGAATGAACTGCCGGGCTATCTGATGGGTTAGA		
	ORF Start: ATG at 21		ORF Stop: TAG at 468
	SEQ ID NO: 64	149 aa	MW at 16975.3kD
NOV19a, CG140188-01 Protein Sequence	METLYRVPFLVLECPNLKPKPPWLQVLSAMIVYALMVVSYFLVTGGIIYDVIVEPPSIG SMTDEHGHQRPVAFLAYRVNEQCIMEGLASSFLFTIGGLGFIFLDRWNAPNIPKLNRFLL LFIFGVCVLLSFFMARVFMRMKLPGYLMG		

Further analysis of the NOV19a protein yielded the following properties shown in
 5 Table 19B.

Table 19B. Protein Sequence Properties NOV19a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.0300 probability located in mitochondrial inner membrane
SignalP analysis:	Cleavage site between residues 48 and 49

A search of the NOV19a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 19C.

Table 19C. Geneseq Results for NOV19a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV19a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAY53631	A bone marrow secreted protein designated BMS155 - <i>Homo sapiens</i> , 149 aa. [WO9933979-A2, 08-JUL-1999]	1..149 1..149	137/149 (91%) 142/149 (94%)	1e-75
AAY53042	Human secreted protein clone pu282_10 protein sequence SEQ ID NO:90 - <i>Homo sapiens</i> , 149 aa. [WO9957132-A1, 11-NOV-1999]	1..149 1..149	137/149 (91%) 142/149 (94%)	1e-75
AAB12143	Hydrophobic domain protein isolated from WERI-RB cells - <i>Homo sapiens</i> , 149 aa. [WO200029448-A2, 25-MAY-2000]	1..149 1..149	137/149 (91%) 142/149 (94%)	1e-75
AAY59670	Secreted protein 108-005-5-0-F6-FL - <i>Homo sapiens</i> , 149 aa. [WO9940189-A2, 12-AUG-1999]	1..149 1..149	137/149 (91%) 142/149 (94%)	1e-75
AAY60146	Human endometrium tumour EST encoded protein 206 - <i>Homo sapiens</i> , 171 aa. [DE19817948-A1, 21-OCT-1999]	1..149 23..171	137/149 (91%) 142/149 (94%)	1e-75

In a BLAST search of public sequence databases, the NOV19a protein was found to have homology to the proteins shown in the BLASTP data in Table 19D.

Table 19D. Public BLASTP Results for NOV19a				
Protein Accession Number	Protein/Organism/Length	NOV19a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9NRP0	DC2 (Hydrophobic protein HSF-28) (Hypothetical 16.8 kDa protein) - <i>Homo sapiens</i> (Human), 149 aa.	1..149 1..149	137/149 (91%) 142/149 (94%)	4e-75
Q9P075	HSPC307 - <i>Homo sapiens</i> (Human), 167 aa (fragment).	1..149 19..167	137/149 (91%) 142/149 (94%)	4e-75

Q9CPZ2	2310008M10Rik protein (RIKEN cDNA 2310008M10 gene) - <i>Mus musculus</i> (Mouse), 149 aa.	1..149 1..149	136/149 (91%) 142/149 (95%)	9e-75
Q9PIR4	HDCMD45P - <i>Homo sapiens</i> (Human), 160 aa (fragment).	1..149 12..160	136/149 (91%) 141/149 (94%)	3e-74
Q8TBU1	Similar to DC2 protein - <i>Homo sapiens</i> (Human), 119 aa.	31..149 1..119	118/119 (99%) 118/119 (99%)	4e-63

PFam analysis predicts that the NOV19a protein contains the domains shown in Table 19E.

Table 19E. Domain Analysis of NOV19a			
Pfam Domain	NOV19a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 20.

The NOV20 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 20A.

Table 20A. NOV20 Sequence Analysis			
	SEQ ID NO: 65	755 bp	
NOV20a, CG140305-01 DNA Sequence	GGAGCTCTGCTGCTCTTCTCAGGGAGACTCTGAGGCTCTGTGAGAATCATGCTTTGGAGG CAGCTCATCTATTGGCAACTGCTGGCTTTGTTTTCTCCCTTTTGGCTGTGTCAAGAT GAATACATGGAGTCTCCACAAACCGGAGGACTACCCCCAGACTGCAGTAAGTGTGTGCAT GGAGACTACAGCTTTCGAGGCTACCAAGCCCCCTGGGCCACCGGGCCCTCCTGGCATT CCAGGAAACCATGGAAACAATGGCAACAATGGAGCCACTGGTCATGAAGGAGCCAAAGGT GAGAAGGGCTACCCGGGGATTCCACCAGAACTTCAGATTGCATTTCATGGCTTCTCTGGCA ACCCACTTCAGCAATCAGAACAGTGGGATTATCTTCAGCAGTGTTGAGACCAACATTGGA AACTTCTTTGATGTCATGACTGGTAGATTGGGGCCCCAGTATCAGGTGTGTATTTCTTC ACCTTCAGCATGATGAAGCATGAGGATGTTGAGGAAGTGTATGTGTACCTTATGCACAAT GGCAACACAGTCTTCAGCATGTACAGCTATGAAATGAAGGGCAAATCAGATACATCCAGC AATCATGCTGTGCTGAAGCTAGCCAAAGGGGATGAGGTTTGGCTGCGAATGGGCAATGGC GCTCTCCATGGGGACCACCAACGCTTCTCCACCTTTGCAGGATTCCTGCTCTTTGAAACT AAGTAAATATATGACTAGTAATAGCTCCACTTTGGG		
	ORF Start: ATG at 49		ORF Stop: TAA at 724
	SEQ ID NO: 66	225 aa	MW at 24836.9kD
NOV20a, CG140305-01 Protein Sequence	MLWRQLIYWQLLALFFLPFCLCQDEYMESPQTGGLPPDCSKCCHGDYSFRGYQGP PGP PPGI PGNHGNNNGNNGATGHEGAKGEKGYPGI PPQLQIAFMASLATHFSNQNSGI IFSSVE TNIGNFFDVMTRFGAPVSGVYFFTFSMMKHEDVEEVYVLMHNGNTVFSMYSYEMKGKS DTSSNHAVLKLAKGDEVWLRMGNGALHGDHQRFS TFAGFLLFETK		
	SEQ ID NO: 67	842 bp	
NOV20b, CG140305-02	GGAGCTCTGCTGCTCTTCTCAGGTAGACTCTGAGGCTCTGTGAGAATCATGCTTTGGAGG CAGCTCATCTATTGGCAACTGCTGGCTTTGTTTTCTCCCTTTTGGCTGTGTCAAGAT		

DNA Sequence	GAATACATGGAGTCTCCACAAACCGGAGGACTACCCCCAGACTGCAGTAAGTGTGTCAT GGAGACTACAGCTTTCGAGGCTACCAAGGCCCTTGGGCCACCGGGCCCTCCTGGCATT CCAGGAAACCATGGAAACAATGGCAACAATGGAGCCACTGGTCATGAAGGAGCCAAAGGT GAGAAGGGCGACAAAGGTGACCTGGGGCCTCGAGGGGAGCGGGGCGAGCATGGCCCCAAA GGAGAGAAGGGCTACCCGGGGATTCCACCAGAACTTCAGATTGCATTTCATGGCTTCTCTG GCAACCCACTTCAGCAATCAGAACAGTGGGATTATCTTCAGCAGTGTTGAGACCAACATT GGAAACTTCTTTGATGTCATGACTGGTAGATTGGGGCCCCAGTATCAGGTGTGTATTTC TTCACCTTCAGCATGATGAAGCATGAGGATGTTGAGGAAGTGTATGTGTACCTTATGCAC AATGGCAACACAGTCTTCAGCATGTACAGCTATGAAATGAAGGGCAAATCAGATACATCC AGCAATCATGCTGTGCTGAAGCTAGCCAAAGGGGATGAGGTTTGGCTGCGAATGGGCAAT GGCGCTCTCCATGGGGACCACCAACGCTTCTCCACCTTTCAGGATTCTGCTCTTTGAA ACTAAGTAAATATATGACTAGTAATAGCTCCACTTTGGGGAAGACTTGTAGCTGAGCTGAT AA		
	ORF Start: ATG at 49		ORF Stop: TAA at 787
	SEQ ID NO: 68	246 aa	MW at 26994.2kD
NOV20b, CG140305-02 Protein Sequence	MLWRQLIYWQLLALFFLPFCLCQDEYMESPQTGGLPPDCSKCCHGDYSFRGYQGP PGPG PPGIPGNHGNNGNNGATGHEGAKGEKGDGDLGPRGERGQHGPKEGKYPGIPPELQIAF MASLATHFSNQNSGIIFSSVETNIGNFFDVMTGRFGAPVSGVYFFTFSMMKHEDVEEVYV YLMHNGNTVFSMSYEMKGKSDTSSNHAVLKLAKGDEVWLRMGNGALHGDHQRFSFAGF LLFETK		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 20B.

Table 20B. Comparison of NOV20a against NOV20b.		
Protein Sequence	NOV20a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV20b	1..225	188/246 (76%)
	1..246	188/246 (76%)

Two polymorphic variants of NOV20a have been identified and are shown in Table 41F. Further analysis of the NOV20a protein yielded the following properties shown in

5 Table 20C.

Table 20C. Protein Sequence Properties NOV20a	
PSort analysis:	0.7666 probability located in outside; 0.2383 probability located in microbody (peroxisome); 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 23 and 24

A search of the NOV20a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 20D.

Table 20D. Geneseq Results for NOV20a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV20a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU84371	Novel human secreted or membrane-associated protein #10 - <i>Homo sapiens</i> , 246 aa. [WO200204600-A2, 17-JAN-2002]	1..225 1..246	225/246 (91%) 225/246 (91%)	e-134
AAB88447	Human membrane or secretory protein clone PSEC0232 - <i>Homo sapiens</i> , 246 aa. [EP1067182-A2, 10-JAN-2001]	1..225 1..246	225/246 (91%) 225/246 (91%)	e-134
AAB18909	A novel polypeptide designated PRO1484 - <i>Homo sapiens</i> , 246 aa. [WO200056889-A2, 28-SEP-2000]	1..225 1..246	225/246 (91%) 225/246 (91%)	e-134
AAB29580	Human adipocyte complement related protein homolog zacrp3, SEQ ID NO:2 - <i>Homo sapiens</i> , 246 aa. [WO200063377-A1, 26-OCT-2000]	1..225 1..246	225/246 (91%) 225/246 (91%)	e-134
AAB15548	Human immune system molecule from Incyte clone 1890540 - <i>Homo sapiens</i> , 246 aa. [WO200060080-A2, 12-OCT-2000]	1..225 1..246	225/246 (91%) 225/246 (91%)	e-134

In a BLAST search of public sequence databases, the NOV20a protein was found to have homology to the proteins shown in the BLASTP data in Table 20E.

Table 20E. Public BLASTP Results for NOV20a				
Protein Accession Number	Protein/Organism/Length	NOV20a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9BXJ4	Complement-clq tumor necrosis factor-related protein 3 precursor (Secretory protein CORS26) - <i>Homo sapiens</i> (Human), 246 aa.	1..225 1..246	225/246 (91%) 225/246 (91%)	e-134

Q9ES30	Collagenous repeat-containing sequence of 26kDa protein - <i>Mus musculus</i> (Mouse), 246 aa.	1..225 1..246	215/246 (87%) 217/246 (87%)	e-127
CAC51163	Sequence 59 from Patent WO0149728 - <i>Homo sapiens</i> (Human), 223 aa.	28..126 101..220	98/120 (81%) 99/120 (81%)	2e-53
Q9ESN4	Gliacolin precursor - <i>Mus musculus</i> (Mouse), 255 aa.	45..222 64..253	66/194 (34%) 97/194 (49%)	1e-22
Q8TE71	EEG1L - <i>Homo sapiens</i> (Human), 1077 aa.	88..223 940..1076	51/138 (36%) 87/138 (62%)	3e-22

Pfam analysis predicts that the NOV20a protein contains the domains shown in Table 20F.

Table 20F. Domain Analysis of NOV20a			
Pfam Domain	NOV20a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Collagen	37..95	23/60 (38%) 37/60 (62%)	0.00032
Clq	98..221	45/137 (33%) 76/137 (55%)	2.3e-17

Example 21.

The NOV21 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 21A.

Table 21A. NOV21 Sequence Analysis			
	SEQ ID NO: 69	1725 bp	
NOV21a, CG140639-01 DNA Sequence	CGGCCGGCGCTGCAGACCCGCTGCTGTTGTCCGGGCTGTGCGGTCCCAGGGCCCTCCG TGCCGCCGGCGCCATGGGCAATTGCCACACGGTGGGGCCCAACGAGGCGTGGTGGTTTC AGGGGGCTGTTGTGGTTCCGACTATAAACAGTACGTGTTTGGCGGCTGGGCCTGGGCCTG GTGGTGTATCTCCGACACTCAGAGGATTTCCCTAGAGATTATGACGTTGCAGCCCCGCTG CGAGGACGTAGAGACGGCCGAGGGGTAGCTTTAACTGTGACGGGTGTCGCCAGGTGAA GATCATGACGGAGAAGGAACCTCTGGCCGTGGCTTGTGAGCAGTTTCTGGGTAAGAATGT GCAGGACATCAAAAACGTCGTCCTGCAGACCCTGGAGGGACATCTGCGCTCCATCCTCGG GACCCTGACAGTGGAGCAGATTTATCAGGACCGGGACCAAGTTGCCAAGCTGGTGCGGGA GGTGGCAGCCCCCTGATGTTGGCCGCATGGGCATTGAGATCCTCAGCTTCACCATCAAGGA CGTGTATGACAAAGTGGACTATCTGAGCTCCCTGGGCAAGACGCAGACTGCCGTGGTGCA GAGAGATGCTGACATTGGCGTGGCCGAGGCTGAACGGGACGCAGGCATCCGGGAAGCTGA GTGCAAGAAGGAGATGCTGGATGTGAAGTTCATGGCAGACACCAAGATTGCTGACTCTAA GCGAGCCTTCGAGCTGCAAAAGTCAGCCTTCAGTGAGGAGGTTAACATCAAGACAGCTGA GGCCAGTTGGCCTATGAGCTGCAGGGGGCCCGTGAACAGCAGAAGATCCGGCAGGAAGA GATTGAGATTGAGGTTGTGCAGCGCAAGAAACAGATTGCCGTGGAGGCACAGGAGATCCT GCGTACGGACAAGGAGCTCATCGCTACAGTGCGCCGGCCTGCCGAGGCCAGGCCACCG CATCCAGCAGATTGCCGAGGGTGAAGAGGTGAAGCAGGTCCTCTTGGCAGGCAGAGGC TGAGAAGATCCGCAAAATCGGGGAGGCGGAAGCGGCAGTCATCGAGGCGATGGGCAAGGC		

	AGAGGCTGAGCGGATGAAGCTCAAGGCAGAAGCCTACCAGAAATACGGGGATGCAGCCAA GATGGCCTTGGTGCTAGAGGCCCTGCCCCAGATTGCTGCCAAAATCGCTGCCCCACTTAC CAAGGTCGATGAGATTGTGGTCTCAGTGGAGACAACAGTAAGGTCACATCAGAAGTGAA CCGACTGCTGGCCGAGCTGCCTGCCTCTGTGCATGCCCTCACAGGCGTGGACCTGTCTAA GATACCCCTGATCAAGAAGGCCACTGGTGTGCAGGTGTGAGGCTCCTGCAGGCCCACTCT CTTCAGCAGCCACCCGGCCCTCCCTCCAGCACCCGTTTAAATCCACAGAACAAACGGGAA CGTTACTGACTCTGGTGCCCTTATCTCGAAGGGACCAGAAGTGCTGCGTGTTCAGGCCATC TCTGGCTGTCTTCCGTCTCTCCTGTCTGTCCACCTCCTCCTCTCCTCTTACCCC ACTTTCAGTGCACCTTCATCAGGTTTGTGTCTCATCTCCCTGCGTGTCTTTCTTTGT CTGTCTTTTCTTTCCCCCATGCACATCATGTAGATTAAAGCTGAAGATGTTTATTACAAT CACTCTCTGTGGGGGTGGCCCTGCTGCTCCTCAGAATCCTGGTG		
	ORF Start: ATG at 74		ORF Stop: TGA at 1358
	SEQ ID NO: 70	428 aa	MW at 47063.7kD
NOV21a, CG140639-01 Protein Sequence	MGNCHTVGPNEALVVS GCCGSDYKQYVFGGWAWAWWCISDTQRISLEIMTLQPRCEDE TAEGVALTVTGVAQVKIMTEKELLAVACEQFLGKNVQDIKNVVLQTLLEGHLRSILGLTLV EQIYQDRDQFAKLVREVAAPDVGRMGIEILSFTIKDVYDKVDYLSLGLKTQTAVVQRDAD IGVAEAERDAGIREAECKKEMLDVKFMADTKIADSKRAFELQSAFSEEVNIKTAEQA YELQGAREQQKIRQEEIEIEVVQRKKQIAVEAQEILRTDKELIATVRRPAEAEAHRIQ AEGEKVKQVLLAQAEAEKIRKIGEAEEAVIEAMGKAEARMKLA EAYQKYGDAAKMALV LEALPQIAAKIAAPLTKVDEIVVLSGDN SKVTSEVNRLLAELPASVHALTGVDLSKIPLI KKATGVQV		
	SEQ ID NO: 71	1389 bp	
NOV21b, CG140639-02 DNA Sequence	CTGCTGTGTCCGGGTCTGTGCGGTCCCGAGGGCCCTCCGTGCCGCCGGCGCCATGGGCA ATTGCCACACGGTGGGGCCCAACGAGGCGCTGGTGGTTTCAGGGGGCTGTGTGGTTCG ACTATAAACAGTACGTGTTTGGCGGCTGGGCCTGGGCCTGGTGGTGTATCTCCGACACTC AGAGGATTTCCCTAGAGATTATGACGTTGCAGCCCCGCTGCGAGGACGTAGAGACGGCCG AGGGGGTAGCTTTAACTGTGACGGGTGTGCGCCAGGTGAAGATCATGACGGAGAAGGAAC TCCTGGCCGTGGCTGTGAGCAGTTTCTGGGTAAGAATGTGCAGGACATCAAAAACGTG TCCTGCAGACCCTGGAGGGACATCTGCGCTCCATCCTCGGGACCCTGACAGTGGAGCAGA TTTATCAGGACCGGGACCAGTTTGCCAAGCTGGTGCGGGAGGTGGCAGCCCTGATGTTG GCCGCATGGGCATTGAGATCCTCAGCTTACCATCAAGGACGTGTATGACAAAGTGGACT ATCTGAGCTCCCTGGGCAAGACGCAGACTGCCGTGGTGCAGAGAGATGCTGACATTGGCG TGGCCGAGGCTGAACGGGACGCAGGCATCCGGGAAGCTGAGTGAAGAAGGAGATGCTGG ATGTGAAGTTTATGGCAGACACCAAGATTGCTGACTCTAAGCGAGCCTTCGAGCTGCAAA AGTCAGCCTTCACTGAGGAGGTTAATCAAGACAGCTGAGGCCAGTTGGCCTATGAGC TGCAGGGGGCCCGTGAACAGCAGAAGATCCGGCAGGAAGAGATTGAGATTGAGTTGTGC AGCGCAAGAAACAGATTGCCGTGGAGGCACAGGAGATCCTGCGTACGGACAAGGAGCTCA TCGCTACAGTGCGCCGGCTGCCGAGGCCGAGGCCACCGCATCCAGCAGATTGCCGAGG GTGAAAAGGTGAAGCAGGTCCTCTTGGCACAGGCAGAGGCTGAGAAGATCCGCAAAATCG GGGAGGCGGAAGCGGCAGTCATCGAGGCGATGGGCAAGGCAGAGGCTGAGCGGATGAAGC TCAAGGCAGAAGCCTACCAGAAATACGGGGATGCAGCCAAGATGGCCTTGGTGCTAGAGG CCCTGCCCCAGATTGCTGCCAAAATCGCTGCCCACTTACCAAGGTGATGAGATTGTGG TCCTCAGTGGAGACAACAGTAAGGTCACATCAGAAGTGAACCGACTGCTGGCCGAGCTGC CTGCCTCTGTGCATGCCCCACAGGCGTGGACCTGTCTAAGATACCCCTGATCAAGAAGG CCACTGGTGTGCAGGTGTGAGGCTCCTGCAGGCCCACTCTCTTCAGCAGGCCACCGGCC TCCCTCCAG		
	ORF Start: ATG at 54		ORF Stop: TGA at 1338
	SEQ ID NO: 72	428 aa	MW at 47047.6kD
NOV21b, CG140639-02 Protein Sequence	MGNCHTVGPNEALVVS GCCGSDYKQYVFGGWAWAWWCISDTQRISLEIMTLQPRCEDE TAEGVALTVTGVAQVKIMTEKELLAVACEQFLGKNVQDIKNVVLQTLLEGHLRSILGLTLV EQIYQDRDQFAKLVREVAAPDVGRMGIEILSFTIKDVYDKVDYLSLGLKTQTAVVQRDAD IGVAEAERDAGIREAECKKEMLDVKFMADTKIADSKRAFELQSAFSEEVNIKTAEQA YELQGAREQQKIRQEEIEIEVVQRKKQIAVEAQEILRTDKELIATVRRPAEAEAHRIQ AEGEKVKQVLLAQAEAEKIRKIGEAEEAVIEAMGKAEARMKLA EAYQKYGDAAKMALV LEALPQIAAKIAAPLTKVDEIVVLSGDN SKVTSEVNRLLAELPASVHAPTGV DLSKIPLI KKATGVQV		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 21B.

Table 21B. Comparison of NOV21a against NOV21b.		
Protein Sequence	NOV21a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV21b	1..428 1..428	407/428 (95%) 407/428 (95%)

Further analysis of the NOV21a protein yielded the following properties shown in Table 21C.

Table 21C. Protein Sequence Properties NOV21a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3000 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

- 5 A search of the NOV21a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 21D.

Table 21D. Geneseq Results for NOV21a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV21a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAW38288	Epidermal surface antigen - <i>Homo sapiens</i> , 379 aa. [US5691460-A, 25-NOV- 1997]	50..428 1..379	377/379 (99%) 377/379 (99%)	0.0
AAR51108	Human epidermal surface antigen - <i>Homo sapiens</i> , 291 aa. [WO9407906-A, 14- APR-1994]	50..326 1..277	276/277 (99%) 276/277 (99%)	e-148
ABB69326	<i>Drosophila melanogaster</i> polypeptide SEQ ID NO 34770 - <i>Drosophila</i> <i>melanogaster</i> , 378 aa. [WO200171042-A2, 27-SEP- 2001]	50..417 1..369	243/370 (65%) 307/370 (82%)	e-134

ABB62956	<i>Drosophila melanogaster</i> polypeptide SEQ ID NO 15660 - <i>Drosophila melanogaster</i> , 426 aa. [WO200171042-A2, 27-SEP-2001]	6..416 7..421	202/417 (48%) 301/417 (71%)	e-104
ABB65943	<i>Drosophila melanogaster</i> polypeptide SEQ ID NO 24621 - <i>Drosophila melanogaster</i> , 430 aa. [WO200171042-A2, 27-SEP-2001]	6..416 7..425	202/421 (47%) 301/421 (70%)	e-102

In a BLAST search of public sequence databases, the NOV21a protein was found to have homology to the proteins shown in the BLASTP data in Table 21E.

Table 21E. Public BLASTP Results for NOV21a				
Protein Accession Number	Protein/Organism/Length	NOV21a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9Z2S9	Flotillin-2 (Reggie-1) (REG-1) - <i>Rattus norvegicus</i> (Rat), 428 aa.	1..428 1..428	425/428 (99%) 426/428 (99%)	0.0
Q9DC36	Adult male lung cDNA, RIKEN full-length enriched library, clone:1200003P16, full insert sequence - <i>Mus musculus</i> (Mouse), 428 aa.	1..428 1..428	424/428 (99%) 425/428 (99%)	0.0
Q9BT16	Similar to flotillin 2 - <i>Homo sapiens</i> (Human), 385 aa.	1..375 1..375	374/375 (99%) 374/375 (99%)	0.0
Q14254	Flotillin-2 (Epidermal surface antigen) (ESA) - <i>Homo sapiens</i> (Human), 379 aa.	50..428 1..379	379/379 (100%) 379/379 (100%)	0.0
Q60634	Flotillin-2 (Epidermal surface antigen) (ESA) - <i>Mus musculus</i> (Mouse), 379 aa.	50..428 1..379	376/379 (99%) 377/379 (99%)	0.0

PFam analysis predicts that the NOV21a protein contains the domains shown in Table 21F.

Table 21F. Domain Analysis of NOV21a			
Pfam Domain	NOV21a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Band_7	12..190	37/215 (17%) 99/215 (46%)	0.28
--------	---------	------------------------------	------

Example 22.

The NOV22 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 22A.

Table 22A. NOV22 Sequence Analysis			
	SEQ ID NO: 73	1201 bp	
NOV22a, CG140843-01 DNA Sequence	CCGCGGAGTGACGCGACCGCGCCGCGCTGAGGGAGGCGCCCCACCATGCCGCGGGCCCC GGCGCCGCTGTACGCCTGCCTCCTGGGGCTCTGCGCGCTCCTGCCCCGGCTCGCAGGTCT CAACATATGCACTAGTGGAAGTGCCACCTCATGTGAAGAATGTCTGCTAATCCACCCAAA ATGTGCCTGGTGCTCCAAAGAGGACTTCGGAAGCCCACGGTCCATCACCTCTCGGTGTGA TCTGAGGGCAAACCTTGTCAAAAATGGCTGTGGAGGTGAGATAGAGAGCCAGCCAGCAG CTTCCATGTCCTGAGGAGCCTGCCCCCTCAGCAGCAAGGGTTCGGGCTCTGCAGGCTGGGA CGTCATTGAGATGACACCACAGGAGATTGCCGTGAACCTCCGGCCCCGGTGACAAGACCAC CTTCCAGCTACAGGTTGCCAGGTGGAGGACTATCCTGTGGACCTGTACTACCTGATGGA CCTCTCCCTGTCCATGAAGGATGACTTGGACAATATCCGGAGCCTGGGCACCAAACTCGC GGAGGAGATGAGGAAGCTCACCAGCAACTTCCGGTTGGGATTGGGTCTTTGTTGATAA GGACATCTCTCCTTCTCCTACACGGCACCGAGGTACCAGACCAATCCGTGCATTGGTTA CAAGTTGTTTCCAAATTGCGTCCCCTCCTTTGGGTTCGCCATCTGCTGCCTCTCACAGA CAGAGTGACAGCTTCAATGAGGAAGTTCGGAACAGAGGGTGTCCCGGAACCGAGATGC CCCTGAGGGGGGCTTTGATGCAGTACTCCAGGCAGCCGTCTGCAAGGTAACCTTCCCTTC TGGTCCCTGTCCCTGCATGGGAGGTCAAGGTAGAGAGCGTCAGTGGGTGTTGGTACTTCC TGCAGGAGTCTTTGAGTGCCCCAGCATGTGGCTCCTGACCACTCTGAAGTCAGAGGGTGA GCTCAGTGGAACCTCTGGGAAATCTACAGCAGTCAAATCAGCCGGAGCTCGGGAATGGAT TGGGCTGGTCTGTGTCTCTGTGTGAGGGTGTGGTTGTGTGCAATGGAGTACTGTCTGCTA GAAGACAGCTGTCTGCAATTTATACATTGGCTTTTTGGTTTATTTTCAGGGGAAAAAAGTA AAGGTCAAGTCATAGGCATAGAAGCTTGTAGAGCTTCTGGACCAATTTTGGCAAACCTT A		
	ORF Start: ATG at 47		ORF Stop: TAG at 1079
	SEQ ID NO: 74	344 aa	MW at 37466.6kD
NOV22a, CG140843-01 Protein Sequence	MPRAPAPLYACLLGLCALLPRLAGLNICTSGSATSCBEECLLIHPKCAWCSKEDFGSPRSI TSRCDLRANLVKNGCGGEIESPASSFHVLRSLPLSSKSGSAGWDVIQMPQEIIVNLRP GDKTTFQLQVRQVEDYPVDLYLMDLSLSMKDDLDNIRSLGTLAEEMRKLTSNFRLGFG SFVDKDISPFSYTAPRYQTNPCIGYKLPNCVPSFGFRHLLPLTDRVDSFNEEVRKQKRV RNRDAPEGGFDAVLQAAVCKVTFSLGPVPAWGGQGRERQWVVLVLPAGVFECPSMWLLTTL KSEGELSGTSGKSTAVKSAGAREWIGLVSVSVSGCGCVQWSTVC		

- One polymorphic variant of NOV22a has been identified and is shown in Table 5
- 41G. Further analysis of the NOV22a protein yielded the following properties shown in Table 22B.

Table 22B. Protein Sequence Properties NOV22a	
PSort analysis:	0.4849 probability located in outside; 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in lysosome (lumen)
SignalP analysis:	Cleavage site between residues 25 and 26

A search of the NOV22a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 22C.

Table 22C. Geneseq Results for NOV22a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV22a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU76337	Human anti-dual integrin protein #3 - <i>Homo sapiens</i> , 799 aa. [WO200212501-A2, 14-FEB-2002]	1..260 1..260	260/260 (100%) 260/260 (100%)	e-153
AAW02194	Human integrin beta subunit protein, beta-5 - <i>Homo sapiens</i> , 799 aa. [US5527679-A, 18-JUN-1996]	1..260 1..260	260/260 (100%) 260/260 (100%)	e-153
AAW13573	Mouse beta-3 integrin - Mus sp, 787 aa. [WO9708316-A1, 06-MAR-1997]	5..259 6..257	149/260 (57%) 186/260 (71%)	5e-77
AAW13574	Mouse beta-3 integrin (truncated) - Mus sp, 720 aa. [WO9708316-A1, 06-MAR-1997]	5..259 6..257	149/260 (57%) 186/260 (71%)	5e-77
AAU76336	Human anti-dual integrin protein #2 - <i>Homo sapiens</i> , 788 aa. [WO200212501-A2, 14-FEB-2002]	5..259 7..258	149/260 (57%) 184/260 (70%)	1e-76

- 5 In a BLAST search of public sequence databases, the NOV22a protein was found to have homology to the proteins shown in the BLASTP data in Table 22D.

Table 22D. Public BLASTP Results for NOV22a				
Protein Accession Number	Protein/Organism/Length	NOV22a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
A38308	integrin beta-5 chain precursor - human, 799 aa.	1..260 1..260	260/260 (100%) 260/260 (100%)	e-153

P18084	Integrin beta-5 precursor - <i>Homo sapiens</i> (Human), 799 aa.	1..260 1..260	260/260 (100%) 260/260 (100%)	e-153
O70309	Integrin beta-5 precursor - <i>Mus musculus</i> (Mouse), 798 aa.	1..260 1..260	241/260 (92%) 252/260 (96%)	e-141
Q8SQB9	Integrin beta 5 subunit precursor protein - <i>Bos taurus</i> (Bovine), 800 aa.	1..260 1..260	235/260 (90%) 246/260 (94%)	e-137
Q9GK49	Integrin beta-5 subunit - <i>Bos taurus</i> (Bovine), 791 aa (fragment).	11..260 2..251	225/250 (90%) 235/250 (94%)	e-131

PFam analysis predicts that the NOV22a protein contains the domains shown in Table 22E.

Table 22E. Domain Analysis of NOV22a			
Pfam Domain	NOV22a Match Region	Identities/ Similarities for the Matched Region	Expect Value
integrin_B	35..260	142/230 (62%) 225/230 (98%)	1.4e-185

Example 23.

The NOV23 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 23A.

Table 23A. NOV23 Sequence Analysis			
	SEQ ID NO: 75	1272 bp	
NOV23a, CG141540-01 DNA Sequence	CCTAGGCCACGTGCTGCTGGGTCTCAGTCCTCCACTTCCCGTGTCTCTGGAAGTTGTCA GGAGCAATGTTGCGCTTGACGTGTTGGTAATGGGAGTTTCTGCCTCCACCCTTCAGCCT GCGGCACACACAGGGGCTGCCAGAAGCTGCCGGTTTCGTGGGAGGCATTACAAGCGGGAG TTCAGGCTGGAAGGGGAGCCTGTAGCCCTGAGGTGCCCCAGGTGCCCTACTGGTTGTGG GCCTCTGTGAGCCCCCGCATCAACCTGACATGGCATAAAATGACTCTGCTAGGACGGTC CCAGGAGAAGAAGAGACCGATGTGGGCCAGGACGGTGCTCTGTGGCTTCTGCCAGCC TTGCAGGAGGACTCTGGCACCTACGTCTGCACTACTAGAAATGCTTCTTACTGTGACAAA ATGTCCATTGAGCTCAGAGTTTTTGAGAATACAGATGCTTTCCTGCCGTTTCATCTCATA CCGCAAATTTTAACTTGTCAACCTCTGGGGTATTAGTATGCCCTGACCTGAGTGAATTC ACCCGTGACAAAACCTGACGTGAAGATTCAATGGTACAAGGATTCTCTTCTTTGGATAAA GACAATGAGAAATTTCTAAGTGTGAGGGGGACCACTCACTTACTCGTACACGATGTGGCC CTGGAAGATGCTGGCTATTACCGCTGTGTCTGACATTTGCCCATGAAGGCCAGCAATAC AACATCACTAGGAGTATTGAGCTACGCATCAAGAGGTCAAGACTGACAAATCCCGTGTAA GTGTTCTGGGAACCGGCACACCTTAACCACCATGCTGTGGTGGACGGCAATGACACC CACATAGAGAGCGCTACCCGGGAGGCCGCGTGACCGAGGGGCCACGCCAGGAATATTCA GAAAATAATGAGAACTACATTGAAGTGCCATTGATTTTGATCCTGTCAAGAGAGGAT TTGCACATGGATTTTAAATGTGTTGTCCATAATACCTGAGTTTTCAGACACTACGCACC ACAGTCAAGGAAGCCTCTCCACGTTCTCTGGGGCATTGTGCTGGCCCCACTTTCACTG GCCTTCTTGGTTTGGGGGAATATGGATGCACAGACGGTCAAACACAGAAGTGGAAAA GCAGATGGTCTGACTGTGCTATGGCCTCATCATCAAGACTTCAATCCTATCCCAAGTGA		

	AATAAATGGAATGAAATAATTCAAACACAAACTCCGTACGTCTTCTCTTATGGAAGTGGC TGTGTCCTTTTG		
	ORF Start: ATG at 67		ORF Stop: TGA at 1198
	SEQ ID NO: 76	377 aa	MW at 43181.9kD
NOV23a, CG141540-01 Protein Sequence	MLRLYLVMGVSASTLQPAHTGAARSCRFRGRHYKREFRLEGEPVALRCPQVPYWLWAS VSPRINLTWHKNDARTVPGEEETRMWAQDGLWLLPALQEDSGTYVCTTRNASYCDKMS IELRVFENTDAFLPFISYPQILTLSTSGVLVCPDLSEFTRDKTDVKIQWYKDSLLLDKDN EKFLSVRGTTLLVHDVALEDAGYYRCVLTFAHEGQQYNITRSIELRIKRSRLTIPCKVF LGTGTPLTTLMLWWTANDTHIESAYPGRVTEGPRQEYSENNENYIEVPLIFDPVTREDLH MDFKCVVHNTLSFQTLRRTTVKEASSTFSWGIVLAPLSLAFVLVGGIWMHRRCKHRTGKAD GLTVLWPHHQDFQSYPK		
	SEQ ID NO: 77	1286 bp	
NOV23b, CG141540-02 DNA Sequence	GCCACGTGCTGCTGGGTCTCAGTCCCTCCACTTCCCGTGTCTCTGGAAGTTGTGAGGAGC AATGTTGCGCTTGTACGTGTTGGTAATGGGAGTTCTGCCTTCACCCTTCAGCCTGCGGC ACACACAGGGGCTGCCAGAAGCTGCCGGTTTCGTGGGAGGCATTACAAGCGGGAGTTCAG GCTGGAAGGGGAGCCTGTAGCCCTGAGGTGCCCCAGGTGCCCTACTGGTTGTGGGCCTC TGTCAGCCCCCGCATCAACCTGACATGGCATAAAAATGACTCTGCTAGGACGGTCCCAGG AGAAGAAGAGACACGGATGTGGGCCAGGACGGTGCTCTGTGGCTTCTGCCAGCCTTGCA GGAGGACTCTGGCACCTACGTCTGCACTACTAGAAATGCTTCTTACTGTGACAAAATGTC CATTGAGCTCAGAGTTTTTGAGAATACAGATGCTTTCCTGCCGTTTCTCTCATACCCGCA AATTTTAACCTTGTCAACCTCTGGGGTATTAGTATGCCCTGACCTGAGTGAATTCACCCG TGACAAAACCTGACGTGAAGATTCAATGGTACAAGGATTCTCTTCTTTTGATGAAAGACAA TGAGAAATTTCTAAGTGTGAGGGGGACCACTCACTTACTCGTACAGATGTGGCCCTGGA AGATGCTGGCTATTACCGCTGTGTCTGACATTTGCCCATGAAGGCCAGCAATACACAT CACTAGGAGTATTGAGCTACGCATCAAGAAAAAAGAAGAGACCATTCTGTGATCAT TTCCCCCTCAAGACCATATCAGCTTCTCTGGGGTCAAGACTGACAATCCCGTGTAAAGGT GTTTCTGGGAACCGGCACACCCTTAACCACCATGCTGTGGTGGACGGCCAATGACACCCA CATAGAGAGCGCTACCCGGGAGGCCGCGTGACCGAGGGGCCACGCCAGGAATATTCAGA AAATAATGAGAACTACATTGAAGTGCCATTGATTTTGTATCCTGTACAAAGAGAGGATTT GCACATGGATTTTAAATGTGTGTCCATAATACCTGAGTTTTCAGACACTACGCCACCAC AGTCAAGGAAGCCTCCTCCAGTTCTCTGGGGCATTGTGCTGGCCCCACTTTCATGGC CTTCTTGTTTTGGGGGAATATGGATGCACAGACGGTGCAAACACAGAACTGGAAAGC AGATGGTCTGACTGTGCTATGGCCTCATCATCAAGACTTTCATCCTATCCCAAGTGAAA TAAATGGAATGAAATAATTCAAACAC		
	ORF Start: ATG at 62		ORF Stop: TGA at 1256
	SEQ ID NO: 78	398 aa	MW at 45420.6kD
NOV23b, CG141540-02 Protein Sequence	MLRLYLVMGVSAFTLQPAHTGAARSCRFRGRHYKREFRLEGEPVALRCPQVPYWLWAS VSPRINLTWHKNDARTVPGEEETRMWAQDGLWLLPALQEDSGTYVCTTRNASYCDKMS IELRVFENTDAFLPFISYPQILTLSTSGVLVCPDLSEFTRDKTDVKIQWYKDSLLLDKDN EKFLSVRGTTLLVHDVALEDAGYYRCVLTFAHEGQQYNITRSIELRIKKKKEETIPV I I SPLKTIASLSRLTIPCKVFLGTGTPLTTLMLWWTANDTHIESAYPGRVTEGPRQEYSE NNENYIEVPLIFDPVTREDLHMDFKCVVHNTLSFQTLRRTTVKEASSTFSWGIVLAPLSLA FLVLGGIWMHRRCKHRTGKADGLTVLWPHHQDFQSYPK		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 23B.

Table 23B. Comparison of NOV23a against NOV23b.		
Protein Sequence	NOV23a Residues/ Match Residues	Identities/ Similarities for the Matched Region

NOV23b	1..377	375/398 (94%)
	1..398	376/398 (94%)

Six polymorphic variants of NOV23a have been identified and are shown in Table 41H. Further analysis of the NOV23a protein yielded the following properties shown in Table 23C.

Table 23C. Protein Sequence Properties NOV23a	
PSort analysis:	0.4600 probability located in plasma membrane; 0.2676 probability located in microbody (peroxisome); 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 14 and 15

- A search of the NOV23a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 23D.

Table 23D. Geneseq Results for NOV23a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV23a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB08207	Human type II Interleukin-1 receptor - <i>Homo sapiens</i> , 398 aa. [WO200187328-A2, 22-NOV-2001]	1..377 1..398	375/398 (94%) 376/398 (94%)	0.0
AAE16581	Human interleukin-1 receptor DNAX designation 2 (IL-1RD2) protein - <i>Homo sapiens</i> , 398 aa. [US6326472-B1, 04-DEC-2001]	1..377 1..398	375/398 (94%) 376/398 (94%)	0.0
AAU78089	Human interleukin 1R2 (IL-1R2) protein sequence - <i>Homo sapiens</i> , 398 aa. [WO200211767-A2, 14-FEB-2002]	1..377 1..398	375/398 (94%) 376/398 (94%)	0.0
AAM24185	Human EST encoded protein SEQ ID NO: 1710 - <i>Homo sapiens</i> , 398 aa. [WO200154477-A2, 02-AUG-2001]	1..377 1..398	375/398 (94%) 376/398 (94%)	0.0

AAB37792	Human interleukin-1 receptor, type II precursor - <i>Homo sapiens</i> , 398 aa. [WO200064479-A1, 02-NOV-2000]	1..377 1..398	375/398 (94%) 376/398 (94%)	0.0
----------	--	------------------	--------------------------------	-----

In a BLAST search of public sequence databases, the NOV23a protein was found to have homology to the proteins shown in the BLASTP data in Table 23E.

Table 23E. Public BLASTP Results for NOV23a				
Protein Accession Number	Protein/Organism/Length	NOV23a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P27930	Interleukin-1 receptor, type II precursor (IL-1R-2) (IL-1R-beta) (Antigen CDw121b) - <i>Homo sapiens</i> (Human), 398 aa.	1..377 1..398	375/398 (94%) 376/398 (94%)	0.0
Q29612	Interleukin-1 receptor, type II precursor (IL-1R-2) (IL-1R-beta) - <i>Cercopithecus aethiops</i> (Green monkey) (Grivet), 393 aa.	1..372 1..393	342/393 (87%) 351/393 (89%)	0.0
AAB05878	Soluble type II interleukin-1 receptor - <i>Homo sapiens</i> (Human), 296 aa.	1..275 1..296	273/296 (92%) 274/296 (92%)	e-159
Q9N2H5	Interleukin-1 receptor type II precursor - <i>Equus caballus</i> (Horse), 403 aa.	4..376 4..396	258/394 (65%) 297/394 (74%)	e-147
P43303	Interleukin-1 receptor, type II precursor (IL-1R-2) - <i>Rattus norvegicus</i> (Rat), 416 aa.	1..376 1..409	232/411 (56%) 282/411 (68%)	e-127

PFam analysis predicts that the NOV23a protein contains the domains shown in

5 Table 23F.

Table 23F. Domain Analysis of NOV23a			
Pfam Domain	NOV23a Match Region	Identities/ Similarities for the Matched Region	Expect Value
ig	43..110	13/70 (19%) 46/70 (66%)	0.00014

ig	165..209	9/47 (19%) 35/47 (74%)	0.0011
ig	230..307	14/78 (18%) 56/78 (72%)	4.3e-05

Example 24.

The NOV24 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 24A.

Table 24A. NOV24 Sequence Analysis			
	SEQ ID NO: 79	4744 bp	
NOV24a, CG141580-01 DNA Sequence	GCTCGGAAC TACACTTCCC GGCAGAACGCGGGCGCGCAGCGCACC GGGGCTCAGCC ATGGCGACCGTGCTGTCCAGGGCGCTCAAGCTGCCGGGGAAGAAGAGCCAGACCTAGGG GAGTATGATCCACTTACCCAGGCTGACAGTGATGAGAGCGAAGACGATCTGGTGCTTAAC CTGCAGAAGAATGGAGGGGTCAAAAATGGGAAGAGTCCTTTGGGAGAAGCGCCAGAACCC GACTCAGATGCTGAGGTTGCAGAGGCTGCAAGCCACATCTTTCAGAAGTCACCACGGAG GGCTACCCCTCAGAACCCCTTGGGGGCTGGAACAGAAGGCGGCCTCCTCCCTGGTGTC TATGTGCGCACGTCTGTCTTCTGCTGACTTTGGGGATCTCGATGATCCTGGTGCTCCTG TGTGCTTTCCTGATCCCCTGTCTCCAGAGATCTGCACAGCACCTGGAGCCGCCACTTG GGCTCCCAGGGAGGTGGGGACCTGTCTCCATTGGAATTGGCTGATGTGAATGGAGATGGC CTGCGTGATGTGCTTCTCTCTTTGTGATGTCAAGGAACGGGAGTGCAAGTAGGTGTCTCA AGACCAGCTGCTAATCTTGTATGCCTTTCGGGGATGAATGGCAGCACACTGTGGTCTAGT CTTCTCCCTGAGGAGGCTCGAGATATCACATGTTTGGAGCTGATGCCAGGAAGCTTGGCT GAAACCATCTGCCTTGTGACAGGGACACACAAGATGCTCAGCGCATTCAATGCAACGTCA GGGAAAGCCATTTGGACTTTAAACCCAACTACTTGTCCAACGGTACCTTGGCTGCCCCA GTTGTGGTACTGCCAGACTTGGATGAAGACGGTGTTCGAGACCTTGTGGTCTTGGCCATT GGGGAATTGCAGCCAGATCTGTGCTTCTGCTGGTGTCTGGCCGGACCGGAAATCCAGTG GGTCGACCTGTGAAGTACAACATCGTTGGAGTTGGGAATCTGATTGGTCTCAGGTTTAC ATCACCAAAATGGGGCTGTCTACATCCTGTTTGGCTTTGGAAATATACAAGCTGTGCGCA CTGCGGGACATTTTGTTCAGGCCCAAATCGAGACAGCTCACACCTTCTCTGCAGATA GAAGAGCCAGAATGGGAAAAGCGAAGATCCATCAACCTGTCTGAGCTCATTGATGTTTAC AGTGATGGTGTGAACTACTCCAGATGGTGAAGGCACAGATTCCAAGTGCAGCAACCTT CTGATTACAACCAGACAAAGCCTTGTGCTGCTTCGGGGGCAAAATCTGACACCTTACTGG GCATTGAGACTTCAAGGCCTGCGCAGCCAGCCTACTCCTGGATATTTCACTGATGATCAG ACATTAGACTTCCTTCTGCAGATACAGGATGGAGTTGGGATGAAAAAGATGATGGTTGTG GATGGTGACTCTGGCTCCATTGTTTGGAGTTACCGTGCTCCGTGTACATGAAAGAAACG CCAGCCACCTCAGCAGTTACTTCAAGCCAGAAAGTCTGTCTTCTCTTCTGGGCCGAAGGG CTGTCAGCTGCATCTCCCAATTCGATATCATCTAGGAAGTGAAGCCGCCCAGCCTTAC CACCTTTACCTCCTGCATCCTGCTTCCCCTCCATCCTTCTGGATCTGGCCAACACCACC GGCACAGTGACGGCTTCAAGGTTGGAATTAACGACCTCTGGAAAGATGCCTTTTATGTT ACCAGGACAACAGGGCCAAGCTCCGAAGGCCATCCAGCAGCCCTGGTGGTCAGCAAGCTT AGTCTACGGTGGGCACTAATGGAGGGCCAGATGGCTCAGCTACAGGAGTCCACCCCAAAA ATTGGCCGTGGGGAGCTGCGAAGATTCTCTCTAGGATAAAGTTTGTGAAGCTCCCTAC GAGATCTAATCTGATGGAATCTTCAGTTGCAGAAGAAGTGAACAGAGTGGATACCCCTC TACTCTCCTGTCACTGTAATCAGTTCTATGGAGAGAAGACTTCTTCTCCTCATTAC CACCTCCCTGATGGTTGCAAAGGCTTGGGAAGGCATGTTGGAGTCTTTGACGGCAGCATG ATCTATTTGGCTGGGGCATCTTACCTACCTTTTTCAGTCCCTGCATTAATCCCTCTAGGA ACTCTGCGTGGACCGTTTGGAAATGTGAATCTCTTAAGTATTAAATTTTTTGGTATGTC TAATTTATGAAGTCTTGTCTGGGAAAGCCAGTGAAGTCTATGACTAGGAAACATTTTGTG TACATTTGTGCTGTGTGTGTATATTTAGTGTGTGGTGAAGTTATTTCCAGGTATGT CCTAAGCTTCAGGGATCCAGTTTCTTGTCTTCTGAAATATATCTGGTTTGTGTCAT TTTGAGACTTCCAGATGCCCTACCTCTGATGTTGAGGGCACTTATTTCTCTCTTATTC		

	TTTCCCACCTGTACCTTGGCTACTTCCAAATTGTAGACAGAATGAGAAAGATTTATAGTG GAAGACTGAGTTAGCCATCCAAGCATTTCATCTCTCTTGTTTTATATCCTATTTCCTTA GATTTTCCATCCATGTCTATTAAGTGACCACAAGAATAACTATATTCTATCACAGGGG AGCAAGAGGATGTAGTCTCAGTGACCCATCTCTGACCAAGTCCACATGTTGTGTATATG TGGCTCTGATGGTCTGCCAGTCATGATCTTTTCTGTGGCGACATCAGAAGTGTATGT TTGCATGCTGTCTTCAACTTAGAGGAGAACTGGAAGTCAGGAGCCTTTGATGTCCTTATC CTGCTGTATGTCTCTCTGCATCTTTTCTATAGGGCACCCCTCCTTAGCTCCCCTCACTC TGTTTCTCTTCTATTAGGGATATGTTCTGGACTTTTTCTCTGCTACTTGAGTCCAG GATGCAACCATTGTGCTGCATCTCTTCTTCTGTAGACCTTTGAAGCATTGTATTT TGGGAAAATTCTCTGTAAATACTATAACTTTTATAAATGGTTAAGTTATTTAGAATTAT CTCCAGTGCTTACTTCTCCCTTCTTCTGTATAAATCTGCTACTTCAATTAAGTTCTCCTC TAACTTTTAGGTCATTGTTTATATAGCAGAAAATTCAATGTTAGCGGATGGAAAAGTGC TTCTTGAATAACCTTGATAGGTCATCCCTGAGTGACCTCAGGTTCTCTCTTACCTGGG CTTGTATCTTTTTTTTTTTTTTTTTTTTTTTTGTAGACAGAGTTTGTCTTGTGCGCCAG GCTGGAGTGCAAGTGGCACAATCTCGGCTCACTGCAACCTTCGCCTCCTGGGTTCAAGCGA TTCTCCAGCCTTAGCCTCCAAGTAGCTGGGACTACAGGTGCCCGCTACCATGCCCTGGCT AATTTTTTTTTTGTATTTTAGTAGAGACGGGGTTTACCATGTTGGCCAGGCTGGTCA CGAACTCCTGACCTCAGATAATCCACCTGCTTCTGCCTCCCAAAGTGCTGGGATTACAGG CGTGAGCCACCATGCCCGCTGGGCTGTATCTTTAGCTTGTGTTAGTAAAGGATTCT AGAAAATTATGAAGTCCAGATTCAAAGGGATCTCTGTTAATTACCACTGACAGGCATTA TGACCTAACAGGAGGTTGGTAGCAGTAGATCCAAGCATGCATGTTGCCCTGGCCTGTAGAT TGGCCTTATCAGGTTTCTGGGTGCCTCTGCCTTAAGATCCTGAAGGCAAATTTGTTTCA ACAGTTTGAAGTCATCTGTGGGTCCAGCTTGACTTTGGAGGAATAAGAAGATACTTCTA GAGTATGGGAATGATTCCAGATAATTCTGGGATTGAATCTACTTGAGTTAAGGCCT GGGACCTAATTTGGTTTAGTATAGAATTTGAAGAATTAATTTATAGGCAGCTGAATACCC AAACTTGGGTGGTGGTCTGTGGTTGGCTGAGCTGTCCGGGCATAACCTGGTTCTCTG TTATGTTAAGGCTTTCTGGGAAGCCAGCCACTCTGCGCAGGAGTGAAACATGAAGTTGTT TTCTGAGGACCTGTTTGGTGGGATTGTTTGGGCAGAGGACTGTGTTTATGCAGGGCAA TCCCAGAAAGATAAGAGGAAGCTAGAGAACTTAATGTACCTGAATTCTTCATGGTGAT TTGCAAACTAATTAACATAGATTCTTTGACTATGGTAAGTTGAATCTCTCCTTGCCA AACAACTTATAAGTTTAGTTTCTTCTCTCTGACGCCGTACGGAAAGGTGTAAGT GGTGGCTGAAAATTGAGGAAGCTTCATCTGACCAATGTGGGTGCTGGTTCTTGTGAAAT GTGTCCTTAAGCCTCCTTCTCCTTGACGGCAGCCACCCAGGTGTCTAAGATAGGAC ATGCTCCTTTCTTCTCTAATCCCATCCTGAGGTGCGCGCAAAGCCAATATGACCACTA CTGAGAAATAGTAATGACTTCTACAAATGCAAGGGTCTTACCCTCCTCTTCCCTTAAAC ACCCTCCCTTTCTTCTAGACCCGTTTGGCCATCCCCCAAATGTGTGGTGAAGAAC TAATCCCCTGAATGTGAATTGCTATCCTTATTGCCCTATTAAAGAAGAGCCAGCTGGTAT ATTGTCAGGAAGCACTATTAAATGTGAAGTGTATAGAGTAAATAAATAAATACTCTA CAGG		
	ORF Start: ATG at 61		ORF Stop: TAA at 1927
	SEQ ID NO: 80	622 aa	MW at 67037.7kD
NOV24a, CG141580-01 Protein Sequence	MATVLSRALKLPKKSPDLGEYDPLTQADSDSEDDLVNLQKNGGVKNGKSPLEGAPEP DSDAEVAEAAKPHLSEVTTEGYPSEPLGGLEQKAASSLVSVRTSVFLLTLGISIMILVLL CAFLIPCPRDLHSTWSRHLGSQGGDLSPLELADVNGDGLRDVLLSFVMSRNGSAVGVS RPAANLVCLSGMNGSTLWSSLLPEEARDITCLELMPGSLAETICLVGTGTHKMLSAFNATS GKAIWTLNPNYLSNGTLAAPVVLPDLDEDGVRDLVVLAIGELQPDLCFLLVSGRTGNPV GRPVKYNIVGVGNLIGPQVYITNGAVYILFGFGNIQAVALRDI FVQAQNRDSSPPSLQI EEPEWEKRRSINLSELIDVSDGVELLQMVKAPDSNCSNLLITRQSLVLLRGQNLTPYW ALRLQGLRSQPTPGYFTDDQTLDFLLQIQDGVGMKMMVVDGDSGSI VWSYRAPCHMKET PATSAVTSQKSVFLFWAEGLSAASPNSDI ILGTEPPSLHHLVLLHPAFPSILLDLANTT GTVTASEVGINDLWKDAFYVTRTTGPSSEGHPAALVVSKLSLRWALMEGQMAQLQESTPK IGRGELRRFLSRIKFVEAPYEI		

Two polymorphic variants of NOV24a have been identified and are shown in Table 41I. Further analysis of the NOV24a protein yielded the following properties shown in Table 24B.

Table 24B. Protein Sequence Properties NOV24a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.3000 probability located in microbody (peroxisome)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV24a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 24C.

Table 24C. Geneseq Results for NOV24a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV24a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB04610	Human quinoprotein dehydrogenase 33 protein SEQ ID NO:2 - <i>Homo sapiens</i> , 302 aa. [CN1307126-A, 08-AUG-2001]	1..284 1..284	283/284 (99%) 283/284 (99%)	e-160
ABB05665	Human transmembrane protein clone amy2_11d2 #2 - <i>Homo sapiens</i> , 552 aa. [WO200198454-A2, 27-DEC-2001]	61..615 6..548	146/565 (25%) 261/565 (45%)	3e-46
ABB89951	Human polypeptide SEQ ID NO 2327 - <i>Homo sapiens</i> , 552 aa. [WO200190304-A2, 29-NOV-2001]	61..615 6..548	145/565 (25%) 260/565 (45%)	1e-45
ABB89787	Human polypeptide SEQ ID NO 2163 - <i>Homo sapiens</i> , 121 aa. [WO200190304-A2, 29-NOV-2001]	232..324 1..99	83/99 (83%) 87/99 (87%)	3e-39
ABB62154	<i>Drosophila melanogaster</i> polypeptide SEQ ID NO 13254 - <i>Drosophila melanogaster</i> , 989 aa. [WO200171042-A2, 27-SEP-2001]	125..465 153..502	86/378 (22%) 145/378 (37%)	5e-14

5 In a BLAST search of public sequence databases, the NOV24a protein was found to have homology to the proteins shown in the BLASTP data in Table 24D.

Table 24D. Public BLASTP Results for NOV24a				
Protein Accession Number	Protein/Organism/Length	NOV24a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9CXB0	8430419L09Rik protein - <i>Mus musculus</i> (Mouse), 624 aa.	1..622 1..624	544/624 (87%) 580/624 (92%)	0.0
Q9P261	KIAA1467 protein - <i>Homo sapiens</i> (Human), 432 aa (fragment).	191..622 1..432	432/432 (100%) 432/432 (100%)	0.0
Q99L10	Similar to RIKEN cDNA 8430419L09 gene - <i>Mus musculus</i> (Mouse), 183 aa (fragment).	440..622 1..183	152/183 (83%) 164/183 (89%)	5e-84
Q96S30	Hypothetical 69.3 kDa protein - <i>Homo sapiens</i> (Human), 636 aa.	61..615 72..605	145/558 (25%) 261/558 (45%)	1e-46
Q9H0X4	Hypothetical 59.7 kDa protein - <i>Homo sapiens</i> (Human), 552 aa.	61..615 6..548	146/565 (25%) 261/565 (45%)	8e-46

PFam analysis predicts that the NOV24a protein contains the domains shown in Table 24E.

Table 24E. Domain Analysis of NOV24a			
Pfam Domain	NOV24a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 25.

- 5 The NOV25 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 25A.

Table 25A. NOV25 Sequence Analysis			
	SEQ ID NO: 81	905 bp	
NOV25a, CG141643-01 DNA Sequence	AACAGCGGCCCTGCGGCTGGCGCGGGACGGGATGAGGCGCTGCAGTCTCTGCGCTTTC GACGCCGCCCGGGGGCCAGGCGGCTGATGCGTGTGGGCCTCGCGTGATCTTGGTGGGC CACGTGAACCTGCTGCTGGGGCCGCTGCTGCATGGCACCGTCTGCGGCACGTGGCCAAT CCCCGCGGCGCTGTACGCCGGAGTACACCGTAGCCAATGTATCTCTGTCTGGCTCGGGG CTGCTGGTGAGCGCGGCAGGCGACCCGGGCGGGGCGGGGCTCCCGGAGAGCCCAGCAGG CCAAAGGCTTTGTGTCTTCCACAGAGCGTTTCCGTGGGACTTGTGGCCCTCTGGCGTCC		

	AGGAACCTTCTTCGCCCTCCACTGCACTGGGTCTGCTGGCACTAGCTCTGGTGAACCTG CTCTTGTCCGTTGCCTGCTCCCTGGGCTCCTTCTTGCTGTGCTCACTCACTGTGGCCAAC GGTGGCCGCCGCTTATTGCTGACTGCCACCCAGGACTGCTGGATCCTCTGGTACCACTG GATGAGGGGCCGGGACATACTGACTGCCCTTTGACCCCAACAAGAATCTATGATACAGCC TTGGCTCTCTGGATCCCTTCTTTGCTCATGTCTGCAGGGGAGGCTGCTCTATCTGGTTAC TGCTGTGTGGCTGCACTCACTCTACGTGGAGTTGGGCCCTGCAGGAAGGACGGACTTCAG GGGCAGGTAGTAGCTGGGTGTGACGCAAGAGTGAAACAGAAAGCCTGGCAGCCACGGTTT CCTGGGATTAAAGTCAAAGCATTATGAATATGGCACTAAAGTGAAGTGAAGTACCAGACCA ATGATCCTGTAAGGCAGCCACAGAACTAAAAACAACAATTATTATTAACTGCTCTGGA TTCTC		
	ORF Start: ATG at 34		ORF Stop: TGA at 805
	SEQ ID NO: 82	257 aa	MW at 26717.2kD
NOV25a, CG141643-01 Protein Sequence	MRRCSLCAFDAARGPRLMRVGLALILVGHVNLLGAVLHGTVLRHVANPRGAVTPEYTV ANVISVGSGLLVSAAGDPGGGRAPGEPSPKALCLPQSVSVGLVALLASRNLLRPPLHWV LLALALVNLLSVACSLGLLAVSLTVANGRRLIADCHPGLDPLVPLDEGPGHTDCPF DPTRIYDTALALWIPSLMSAGEAALSGYCCVAALTLRGVGPCKDKGLQGQVVGCDARV KQKAWQPRFPFGIKVKAL		

Further analysis of the NOV25a protein yielded the following properties shown in Table 25B.

Table 25B. Protein Sequence Properties NOV25a	
PSort analysis:	0.6400 probability located in plasma membrane; 0.4600 probability located in Golgi body; 0.3700 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 37 and 38

A search of the NOV25a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
5 homologous proteins shown in Table 25C.

Table 25C. Geneseq Results for NOV25a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV25a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AA Y78805	Hydrophobic domain containing protein clone HP10508 protein sequence - <i>Homo sapiens</i> , 231 aa. [WO200000506-A2, 06-JAN-2000]	1..257 1..231	231/257 (89%) 231/257 (89%)	e-127
ABB90256	Human polypeptide SEQ ID NO 2632 - <i>Homo sapiens</i> , 240 aa. [WO200190304-A2, 29-NOV-2001]	1..232 1..206	205/232 (88%) 206/232 (88%)	e-111

AAU83615	Human PRO protein, Seq ID No 48 - <i>Homo sapiens</i> , 222 aa. [WO200208288-A2, 31-JAN-2002]	19..232 1..188	187/214 (87%) 188/214 (87%)	1e-99
AAG81326	Human AFP protein sequence SEQ ID NO:170 - <i>Homo sapiens</i> , 222 aa. [WO200129221-A2, 26-APR-2001]	19..232 1..188	187/214 (87%) 188/214 (87%)	1e-99
AAB43588	Human cancer associated protein sequence SEQ ID NO:1033 - <i>Homo sapiens</i> , 243 aa. [WO200055350-A1, 21-SEP-2000]	102..232 79..209	127/131 (96%) 129/131 (97%)	9e-70

In a BLAST search of public sequence databases, the NOV25a protein was found to have homology to the proteins shown in the BLASTP data in Table 25D.

Table 25D. Public BLASTP Results for NOV25a				
Protein Accession Number	Protein/Organism/Length	NOV25a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
AAH27812	Similar to RIKEN cDNA 2010001C09 gene - <i>Mus musculus</i> (Mouse), 249 aa.	1..232 10..215	185/232 (79%) 194/232 (82%)	e-100
CAC38576	Sequence 169 from Patent WO0129221 - <i>Homo sapiens</i> (Human), 222 aa.	19..232 1..188	187/214 (87%) 188/214 (87%)	4e-99
Q9D817	2010001C09Rik protein - <i>Mus musculus</i> (Mouse), 223 aa.	1..232 10..189	163/232 (70%) 171/232 (73%)	3e-82
Q969K7	Hypothetical 23.8 kDa protein (Similar to RIKEN cDNA 1810017F10 gene) (Beta-casein-like protein) - <i>Homo sapiens</i> (Human), 222 aa.	18..210 17..177	66/193 (34%) 104/193 (53%)	6e-26
Q8VCL0	RIKEN cDNA 1810017F10 gene - <i>Mus musculus</i> (Mouse), 219 aa.	18..210 17..177	69/195 (35%) 101/195 (51%)	1e-24

PFam analysis predicts that the NOV25a protein contains the domains shown in

5 Table 25E.

Table 25E. Domain Analysis of NOV25a			
Pfam Domain	NOV25a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 26.

The NOV26 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 26A.

Table 26A. NOV26 Sequence Analysis			
	SEQ ID NO: 83	446 bp	
NOV26a, CG142003-01 DNA Sequence	CTGGGGATAGAGCCTCCTCAAATCCAAATGCTACCAGCTCCAGCTCCCAGGATCCAGAGA GTTTGCAAGACAGAGGCGAAGGGAAGGTCGCAACAACAGTTATCTCCAAGATGCTATTCG TTGAACCCATCCTGGAGGTTTCCAGCTTGCCGACAACCAACTCAACAACCAATTTCAGCCA CCAAATAACAGCTAATACCACTGATGAACCCACCACACAACCCACCACAGAGGACCCAG ATCTTCAGGTTTCTGCGATGCAGCACCAGACAGTCTGGAAGTACAGAGACTGGGGTGG AGGTGGCTGCAGCCTCCGCCATCTCTGTGGCCCGCACCTGCTGGTCTTTGAAGTGCAGC AGCCCTTCTCTTCGTGCTCTGGGACCAGCAGCACAAGTCCCTGTCTTCATGGGGCGAG TATATGACCCAGGGCCTGAGACAAG		
	ORF Start: at 3		ORF Stop: TGA at 438
	SEQ ID NO: 84	145 aa	MW at 15697.3kD
NOV26a, CG142003-01 Protein Sequence	GDRASSNPATSSSSQDPESLQDRGEGKVATTVISKMLFVEPILEVSSLPTTNSTTNSAT KITANTTDEPTTQPTTEDPDLQVSAMQHQTVLELTETGVEVAAASAI SVARTLLVFEVQQ PFLFVLWDQQHKFPVFMGRVYDPR		
	SEQ ID NO: 85	436 bp	
NOV26b, 306076006 DNA Sequence	CACCAAGCTTAATCCAAATGCTACCAGCTCCAGCTCCCAGGATCCAGAGAGTTTGCAAGA CAGAGGCGAAGGGAAGGTCGCAACAACAGTTATCTCCAAGATGCTATTCTGTTGAACCCAT CCTGGAGGTTTCCAGCTTGCCGACAACCAACTCAACAACCAATTTCAGCCACCAAATAAC AGCTAATACCACTGATGAACCCACCACACAACCCACCACAGAGGACCCAGATCTTCAGGT TTCTGCGATGCAGCACCAGACAGTGTGGAAGTACAGAGACTGGGGTGGAGGTGGCTGC AGCCTCCGCCATCTCTGTGGCCCGCACCTGCTGGTCTTTGAAGTGCAGCAGCCCTTCCT CTTCGTGCTCTGGGACCAGCAGCACAAGTCCCTGTCTTCATGGGGCGAGTATATGACCC CAGGGCCCTCGAGGGC		
	ORF Start: at 2		ORF Stop: end of sequence
	SEQ ID NO: 86	145 aa	MW at 15765.5kD
NOV26b, 306076006 Protein Sequence	TKLNPATSSSSQDPESLQDRGEGKVATTVISKMLFVEPILEVSSLPTTNSTTNSATKIT ANTTDEPTTQPTTEDPDLQVSAMQHQTVLELTETGVEVAAASAI SVARTLLVFEVQQPFL FVLWDQQHKFPVFMGRVYDPR		
	SEQ ID NO: 87	223 bp	
NOV26c, 278889088 DNA Sequence	CACCAAGCTTACAGAGGACCCAGATCTTCAGGTTTCTGCGATGCAGCACCAGACAGTGTCT GGAAGTACAGAGACTGGGGTGGAGGTGGCTGCAGCCTCCGCCATCTCTGTGGCCCGCAC CCTGCTGGTCTTTGAAGTGCAGCAGCCCTTCCTCTTCGTGCTCTGGGACCAGCAGCACA GTTCCCTGTCTTCATGGGGCGAGTATATGACCCCTCGAGGGC		
	ORF Start: at 2		ORF Stop: end of sequence
	SEQ ID NO: 88	74 aa	MW at 8317.5kD
NOV26c, 278889088 Protein Sequence	TKLTEDPDLQVSAMQHQTVLELTETGVEVAAASAI SVARTLLVFEVQQPFLFVLWDQQHK FPVFMGRVYDPR		

278889088			
Protein Sequence			
	SEQ ID NO: 89	529 bp	
NOV26d, CG142003-02 DNA Sequence	GAGGAGAAGTTTGGAGTCCGCTGACGTGCGCGCCAGATGGCCTCCAGGCTGACCCTGCT GACCCTCCTGCTGCTGCTGCTGGCTGGGGATAGAGCCTCCTCAAATCCAAATGCTACCAG CTCCAGCTCCCAGGATCCAGAGAGTTTGCAAGACAGAGGCGAAGGGAAGGTCGCAACAAC AGTTATCTCCAAGATGCTATTCGTTGAACCCATCCTGGAGGTTTCCAGCTTGCCGACAAC CAACTCAACAACCAATTCAGCCACCAAAATAACAGCTAATACCACTGATGAACCCACCAC ACAACCCACCACAGAGGACCAGATCTTCAGGTTTCTGCGATGCAGCACCAGACAGTGCT GGAAGTACAGAGACTGGGGTGGAGGTGGCTGCAGCCTCCGCCATCTCTGTGGCCCGCAC CCTGCTGGTCTTTGAAGTGCAGCAGCCCTTCTCTCTGCTGCTGGGACCAGCAGCACAA GTTCCCTGTCTTCATGGGGCGAGTATATGACCCAGGGCCTGAGACAAG		
	ORF Start: ATG at 38		ORF Stop: TGA at 521
	SEQ ID NO: 90	161 aa	MW at 17434.5kD
NOV26d, CG142003-02 Protein Sequence	MASRLTLLTLLLLLAGDRASSNPATSSSSQDPESLQDRGEGKVATTVISKMLFVEPIL EVSSLPTTNTNSATKITANTTDEPTTQPTTEDPDLQVSAMQHQTVLELTETGVEVAAA SAISVARTLLVFEVQQPFLFVLWDQHKFPVFMGRVYDPRA		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 26B.

Table 26B. Comparison of NOV26a against NOV26b through NOV26d.		
Protein Sequence	NOV26a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV26b	7..145 4..142	99/139 (71%) 99/139 (71%)
NOV26c	76..143 4..71	58/68 (85%) 58/68 (85%)
NOV26d	1..145 17..161	93/145 (64%) 93/145 (64%)

One polymorphic variant of NOV26a has been identified and is shown in Table 41J.

- 5 Further analysis of the NOV26a protein yielded the following properties shown in Table 26C.

Table 26C. Protein Sequence Properties NOV26a	
PSort analysis:	0.6500 probability located in cytoplasm; 0.1555 probability located in lysosome (lumen); 0.1000 probability located in mitochondrial matrix space; 0.0000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV26a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 26D.

Table 26D. Geneseq Results for NOV26a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV26a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU02972	Angiotensin converting enzyme (ACEV) splice variant protein #72 - <i>Homo sapiens</i> , 636 aa. [WO200136632-A2, 25-MAY-2001]	1..94 17..109	81/94 (86%) 83/94 (88%)	3e-37
AAW18207	Wild-type C1 inhibitor - <i>Homo sapiens</i> , 500 aa. [US5622930-A, 22-APR-1997]	1..94 17..109	81/94 (86%) 83/94 (88%)	3e-37
AAW18212	Recombinant C1 inhibitor mutein - <i>Homo sapiens</i> , 500 aa. [US5622930-A, 22-APR-1997]	1..94 17..109	81/94 (86%) 83/94 (88%)	3e-37
AAW18218	Recombinant C1 inhibitor mutein - <i>Homo sapiens</i> , 500 aa. [US5622930-A, 22-APR-1997]	1..94 17..109	81/94 (86%) 83/94 (88%)	3e-37
AAW18217	Recombinant C1 inhibitor mutein - <i>Homo sapiens</i> , 500 aa. [US5622930-A, 22-APR-1997]	1..94 17..109	81/94 (86%) 83/94 (88%)	3e-37

In a BLAST search of public sequence databases, the NOV26a protein was found to have homology to the proteins shown in the BLASTP data in Table 26E.

Table 26E. Public BLASTP Results for NOV26a				
Protein Accession Number	Protein/Organism/Length	NOV26a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96FE0	Serine (or cysteine) proteinase inhibitor, clade G (C1 inhibitor), member 1 - <i>Homo sapiens</i> (Human), 500 aa.	1..94 17..109	81/94 (86%) 83/94 (88%)	8e-37

P05155	Plasma protease C1 inhibitor precursor (C1 Inh) (C1Inh) - <i>Homo sapiens</i> (Human), 500 aa.	1..94 17..109	81/94 (86%) 83/94 (88%)	8e-37
Q95J12	Complement C1 inhibitor - Pan troglodytes (Chimpanzee), 162 aa (fragment).	2..82 1..80	75/81 (92%) 77/81 (94%)	3e-34
Q16304	C1-inhibitor - <i>Homo sapiens</i> (Human), 83 aa (fragment).	76..145 14..83	67/70 (95%) 68/70 (96%)	7e-32
P97290	Plasma protease C1 inhibitor precursor (C1 Inh) (C1Inh) - <i>Mus musculus</i> (Mouse), 504 aa.	76..144 435..503	57/69 (82%) 65/69 (93%)	2e-27

PFam analysis predicts that the NOV26a protein contains the domains shown in Table 26F.

Table 26F. Domain Analysis of NOV26a			
Pfam Domain	NOV26a Match Region	Identities/ Similarities for the Matched Region	Expect Value
serpin	76..143	31/74 (42%) 61/74 (82%)	2.5e-25

Example 27.

The NOV27 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 27A.

Table 27A. NOV27 Sequence Analysis			
	SEQ ID NO: 91	1356 bp	
NOV27a, CG142023-01 DNA Sequence	GGCGAGGCCGGCGCGATGCGGCAGCTGTGCCGGGGCCGCGTGTGGGCATCTCGGTGGCC ATCGCGCACGGGGTCTTCTCGGGCTCCCTCAACATCTTGCTCAAGTTCCTCATCAGCCGC TACCAGTTCTCCTTCTGACCCTGGTGACGTGCCGTGACCAGCTCCACCGCGGCGTGAGC CTGGAGCTGCTGCGGCGCCTCGGGCTCATCGCCGTGCCCCCTTCGGTCTGAGCCTGGCG CGCTCCTTCGCGGGGGTCGCGGTGCTCTCCACGCTGCAGTCCAGCCTCACGCTCTGGTCC CTGCGCGGCCTCAGCCTGCCCATGTACGTGGTCTTCAAGCGCTGCCTGCCCTGGTACC ATGCTCATCGGCGTCCTGGTGCTCAAGAACGGCGCGCCCTCGCCAGGGGTGCTGGCGCG GTGCTCATCACCACTGCGGCGCGCCCTGGCAGGTGCCGGCGACCTGACGGGCGACCCC ATCGGGTACGTACGGGAGTGCTGGCGGTGCTGGTGACGCTGCCTACCTGGTGCTCATC CAGAAGGCCAGCGCAGACACCGAGCACGGGCGGCTCACCAGCGAGTACGTACATCGCCGTC TCTGCCACCCGCTGCTGGTCATCTGCTCCTTCGCCAGCACCGACTCCATCCACGCTGG ACCTTCCCGGGCTGGAAGGACCCGGCCATGGTCTGCATCTTCGTGGCCTGCATCCTGATC GGTGCGCCATGAACCTTACCACGCTGCACTGCACCTACATCAATTTCGGCCGTGACCACC AGCTTCGTGGGTGTGGTGAAGAGCATCGCCACCATCACGGTGGGCATGGTGGCCTTCAGC GACGTGGAGCCACCTCTCTGTTCATTGCCGGCGTGGTGGTGAACACCCCTGGGCTCTATC ATTTACTGTGTGGCCAAGTTCATGGAGACCAGAAAGCAAAGCAACTACGAGGACCTGGAG GCCCAGCCTCGGGGAGAGGAGGCGCAGCTAAGTGGAGACCAGCTGCCGTTCTGTATGGAG GAGCTGCCCGGGGAGGGAGGAAATGCCCGGTACAGAGGTGGGGAGGCAGAGGTGGCCCC		

	GCTCAGGAGAGCAGGCAAGAGGTGAGGGGAGCCCCGAGGAGTCCCCTGGTGGCTGGG AGCTCTGAAGAAGGGAGCAGGAGGTCGTTAAAAGATGCTTACCTCGAGGTATGGAGGTTG GTTAGGGGAACAGGTATATGAAGAAGGATTATTGATAGAAAACGAGGAGTTACCCAGT CCTTGAGAAGGAGGTGCATGTACGTACCTATGTGCATACACTATTTTATATGTTAGAAA TGACGTGTTTTAATGAGAGGCCTCCCCGTTTTATTC		
	ORF Start: ATG at 16		ORF Stop: TGA at 1264
	SEQ ID NO: 92	416 aa	MW at 44181.9kD
NOV27a, CG142023-01 Protein Sequence	MRQLCRGRVLGISVAIAHGVFSGSLNILLKFLISRYQFSFTLVQCLTSSTAALSLELLR RLGLIAVPPFGLSLARFAGVAVLSTLQSSLTLWSLRGLSLPMYVVFKRCLPLVTMLIGV LVLKNGAPSPGVLA AVLITTCGAALAGAGDLTGDPFGYVTGVLAVLVHAYLVLIQKASA DTEHGPLTAQYVIAVSATPLLVICFASTDSIHAWTFPGWKDPAMVCIFVACILIGCAMN FTTLHCTYINSAVTTSFVGVVKSIAITITVGMVAFSDVEPTSLFIAGVVVNTLGSIIYCVA KFMETRKQSNYEDLEAQPRGEEAQLSGDQLPFVMEELPGEGGNGRSEGGEAAGGPAQESR QEVGRSPRGVPLVAGSSEEGSRRLKDAYLEVWRLVRGTRYMKDYLIENEELPSP		

Further analysis of the NOV27a protein yielded the following properties shown in Table 27B.

Table 27B. Protein Sequence Properties NOV27a	
PSort analysis:	0.6400 probability located in plasma membrane; 0.4600 probability located in Golgi body; 0.3700 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 20 and 21

A search of the NOV27a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
5 homologous proteins shown in Table 27C.

Table 27C. Geneseq Results for NOV27a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV27a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU81226	Human lung cancer protein, Seq ID No 62 - <i>Homo sapiens</i> , 391 aa. [WO200192525-A2, 06-DEC-2001]	1..416 1..391	391/416 (93%) 391/416 (93%)	0.0
AAM47572	Drosophila cell cycle progression protein #1 - <i>Drosophila</i> sp, 373 aa. [WO200172774-A2, 04-OCT-2001]	12..321 64..371	87/316 (27%) 153/316 (47%)	3e-21

ABB60236	<i>Drosophila melanogaster</i> polypeptide SEQ ID NO 7500 - <i>Drosophila melanogaster</i> , 373 aa. [WO200171042-A2, 27-SEP-2001]	12..321 64..371	87/316 (27%) 153/316 (47%)	3e-21
AAB88597	Human hydrophobic domain containing protein clone HP03670 #121 - <i>Homo sapiens</i> , 337 aa. [WO200112660-A2, 22-FEB-2001]	8..322 24..329	74/315 (23%) 137/315 (43%)	7e-14
AAB56473	Human prostate cancer antigen protein sequence SEQ ID NO:1051 - <i>Homo sapiens</i> , 341 aa. [WO200055174-A1, 21-SEP-2000]	8..322 28..333	74/315 (23%) 136/315 (42%)	1e-13

In a BLAST search of public sequence databases, the NOV27a protein was found to have homology to the proteins shown in the BLASTP data in Table 27D.

Table 27D. Public BLASTP Results for NOV27a				
Protein Accession Number	Protein/Organism/Length	NOV27a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9CXD4	6230421J19Rik protein - <i>Mus musculus</i> (Mouse), 152 aa.	271..416 1..152	111/152 (73%) 120/152 (78%)	4e-55
Q94B65	Hypothetical 34.6 kDa protein - <i>Arabidopsis thaliana</i> (Mouse-ear cress), 323 aa.	10..319 13..323	93/316 (29%) 163/316 (51%)	8e-34
Q9SB76	Hypothetical 31.9 kDa protein - <i>Arabidopsis thaliana</i> (Mouse-ear cress), 296 aa.	30..319 6..296	90/296 (30%) 151/296 (50%)	1e-31
Q95YI5	UDP-sugar transporter UST74c (Fringe connection protein) - <i>Drosophila melanogaster</i> (Fruit fly), 373 aa.	12..321 64..371	87/316 (27%) 153/316 (47%)	9e-21

Q9NTN3	UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter (UDP-GlcA/UDP-GalNAc transporter) - <i>Homo sapiens</i> (Human), 355 aa.	18..309 49..341	80/295 (27%) 132/295 (44%)	1e-16
--------	--	--------------------	-------------------------------	-------

PFam analysis predicts that the NOV27a protein contains the domains shown in Table 27E.

Table 27E. Domain Analysis of NOV27a			
Pfam Domain	NOV27a Match Region	Identities/ Similarities for the Matched Region	Expect Value
DUF6	166..299	21/135 (16%) 87/135 (64%)	0.29

Example 28.

The NOV28 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 28A.

Table 28A. NOV28 Sequence Analysis			
	SEQ ID NO: 93	785 bp	
NOV28a, CG142092-01 DNA Sequence	AAAAAGCTCCATCTGGGGCTCTTCATAGAAAAAGGAAAATGGCAGCCTGGCCCTTCTCCAG GCTGTGGAAAGTCTCTGATCCAATTCTTCCAAATGACCTTGATCGTGCTCTGTGGCC TGCTGTTCTTGGCAATTGTGGTCTCCACCCACCTTATCATTGCTGCCCCGATGGATAT TACGTTGACTGAGACACGCTTCAAACTGGAACCTACTCTGAAATACACCTGCCTCCCTGG CTACGTCAGATCCCACTCACTCAGACGCTTACCTGTAATTCTGATGGCGAATGGGTGTA TAACACCTTCTGTATCTACAAACGATGCAGACACCCAGGAGAGTTACGTAATGGGCAAGT AGAGATTAAGACAGATTTATCTTTTGGATCACAAATAGAATTCAGCTGTTCAGAAGGATT TTTCTTAATTGGCTCAACCACTAGTCGTTGTGAAGTCCAAGATAGAGGAGTTGGCTGGGG TCATCCTCTCCCAATGTGAAATTGTCAAGTGTAAAGCCTCCTCCAGACATCAGGAATGG AAGGCACAGCGGTGAAGAAAATTTCTACGCATACGCTTTTCTGTACCTACAGCTGTGA ACAAGTGCTCACAGGCAAAAGACTCATGCAGTGTCTCCAAACCCAGAGGATGTGAAAAT GGCCCTGGAGGTATATAAGCTGTCTCTGGAATTGAACAACTGGAACCTACAGAGAGACAG CGCAAGACAATCCACTTTGGATAAAGAACTATAATTTTCTCAAAGAAGGAGGAAAAGG TGTCT		
	ORF Start: at 2		ORF Stop: TAA at 752
	SEQ ID NO: 94	250 aa	MW at 28139.0kD
NOV28a, CG142092-01 Protein Sequence	KTPSGALHRKRKMAAWPFSRLWKVSDPILFQMTLIAALLPAVLGNCGPPTLSFAAPMDI TLTEFRFKTGTLKYTCPLPGYVRSHSTQTLTCNSDGEWVYNTFCIYKRCHPGELRNGQV EIKTDLSEFGSIEFSCSEGFLLIGSTTSRCEVQDRGVGWGHPQPCEIVKCKPPPDIRNG RHSGEENFYAYGFSVTYSCEQVLTGKRLMQCLPNPEDVKMALEVYKLSLEIEQLELQRDS ARQSTLDKEL		
	SEQ ID NO: 95	972 bp	
NOV28b, CG142092-02 DNA Sequence	AAACTCTGATCTGGGGAGGAACCGAGCTACATAGATCAAGGCAGTTTCTTCTTTGAG AAATATCCAGATATCATCATAGAGTCTTCTGCTCTTCTCAACTACCAAGAAAAACA TCAGCGAAGCAGCAGGCCATGCACCCCCCAAACTCCATCTGGGGCTTTCATAGAAAA AGGAAAATGGCAGCCTGGCCCTTCTCCAGGCTGTGGAAAGTCTCTGATCCAATTCTCTT CAAATGACCTTGATCGTGCTCTGTTGCCTGCTGTTCTTGCAATTGTTGCTCTCCACCC		

	ACTTTATCATTGCTGCCCCGATGGATATTACGTTGACTGAGACACGCTTCAAACTGGA ACTACTCTGAAATACACCTGCCTCCCTGGCTACGTCAGATCCCATTCAACTCAGACGCTT ACCTGTAATTCTGATGGCGAATGGGTGTATAACACCTTCTGTATCTACAAACGATGCAGA CACCCAGGAGAGTTACGTAATGGGCAAGTAGAGATTAAGACAGATTTATCTTTTGATCA CAAATAGAATTGAGCTGTTGAGAAGGCTGTGAACAAGTGCTCACAGGCAAAAGACTCATG CAGTGTCTCCCAAACCCAGAGGATGTGAAAATGGCCCTGGAGGTATATAAGCTGTCTCTG GAAATTGAACAACCTGGAACACAGAGAGACAGCGCAAGACAATCCACTTTGGATAAAGAA CTATAATTTTCTCAAAGAAGGAGGAAAAGGTGTCTTGCTGGCTTGCCCTCTGCAATTC AATACAGATCAGTTTAGCAAATCTACTGTCAATTTGGCAGTGATATTATCATAATAAAT ATCTAGAAATGATAATTTGCTAAAGTTTAGTGCTTTGAGATTGTGAAATTATTAATCATC CTCTGTGGCTCATGTTTTTGTCTTTCAACACACAAAGCACAAATTTTTTTCGATTAA AAATGTATGTAT		
	ORF Start: ATG at 139		ORF Stop: TAA at 724
	SEQ ID NO: 96	195 aa	MW at 21984.2kD
NOV28b, CG142092-02 Protein Sequence	MHPPKTPSGALHRKRKMAAWPFSRLWKVSDPILFQMTLIAALLPAVLGNCPPPTLSFAA PMDITLTETRFKTGTTLKYTCLPGYVRSHSTQTLTNCSDGEWVYNTFCIYKRCRHPGELR NGQVEIKTDLSEFGSQIEFSCSEGCEQVLTGKRLMQCLPNPEDVKMALEVYKLSLEIEQLE LORDSARQSTLDKEL		
	SEQ ID NO: 97	681 bp	
NOV28c, CG142092-03 DNA Sequence	AAAACTCTGATCTGGGGAGGAACAGGACTACATAGATCAAGGCAGTTTCTTCTTTGAG AAACATATCCAGATATCATATAGAGTCTTCTGCTCTTCTCAACTACCAAAGAAAAACA TCAGCGAAGCAGCAGGCCATGCACCCCCCAAACTCCATCTGGGGCTCTTCATAGAAAA AGGAAAATGGCAGCCTGGCCCTTCTCCAGGCTGTGGAAAGTCTCTGATCCAATTCTCTTC CAAATGACCTTGATCGCTGCTCTGTTGCTGCTGTTCTTGGCAATTGTGGTCTCCACCC ACTTTATCATTGCTGCCCCGATGGATATTACGTTGACTGAGACACGCTTCAAACTGGA ACTACTCTGAAATTGAACAACCTGGAACACAGAGAGACAGCGCAAGACAATCCACTTTG GATAAAGAACTATAATTTTCTCAAAGAAGGAGGAAAAGGTGTCTTGCTGGCTTGCCCTC TTGCAATTCAATACAGATCAGTTTAGCAAATCTACTGTCAATTTGGCAGTGATATTCATC ATAATAAATATCTAGAAATGATAATTTGCTAAAGTTTAGTGCTTTGAGATTGTGAAATTA TTAATCATCCTCTGTGTGGCTCATGTTTTGCTTTCAACACACAAAGCACAAATTTTTT TTCGATTAAAAATGTATGTAT		
	ORF Start: ATG at 139		ORF Stop: TAA at 433
	SEQ ID NO: 98	98 aa	MW at 10927.6kD
NOV28c, CG142092-03 Protein Sequence	MHPPKTPSGALHRKRKMAAWPFSRLWKVSDPILFQMTLIAALLPAVLGNCPPPTLSFAA PMDITLTETRFKTGTTLEIEQLELQDSARQSTLDKEL		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 28B.

Table 28B. Comparison of NOV28a against NOV28b and NOV28c.		
Protein Sequence	NOV28a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV28b	1..250 5..195	185/250 (74%) 185/250 (74%)
NOV28c	1..74 5..78	73/74 (98%) 74/74 (99%)

Further analysis of the NOV28a protein yielded the following properties shown in Table 28C.

Table 28C. Protein Sequence Properties NOV28a	
PSort analysis:	0.6500 probability located in plasma membrane; 0.5046 probability located in mitochondrial inner membrane; 0.3752 probability located in microbody (peroxisome); 0.3000 probability located in Golgi body
SignalP analysis:	Cleavage site between residues 45 and 46

A search of the NOV28a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 28D.

Table 28D. Geneseq Results for NOV28a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV28a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAR13490	Human C4 binding protein - <i>Homo sapiens</i> , 581 aa. [WO9111461-A, 08-AUG-1991]	13..218 1..208	190/208 (91%) 193/208 (92%)	e-113
AAB57162	Human prostate cancer antigen protein sequence SEQ ID NO:1740 - <i>Homo sapiens</i> , 110 aa. [WO200055174-A1, 21-SEP-2000]	62..170 1..109	107/109 (98%) 108/109 (98%)	1e-61
AAW39924	Amino acid sequence of a mouse sperm protein designated sp56 - <i>Mus sp</i> , 579 aa. [WO9800440-A1, 08-JAN-1998]	13..204 1..192	103/193 (53%) 132/193 (68%)	2e-57
AAG68150	Codon modified human DAF protein sequence SEQ ID NO:1 - <i>Homo sapiens</i> , 320 aa. [JP2001211882-A, 07-AUG-2001]	32..217 22..212	74/191 (38%) 106/191 (54%)	3e-32
ABB07542	Amino acid sequence of APT2334 - Synthetic, 271 aa. [WO200204638-A1, 17-JAN-2002]	45..217 65..241	68/177 (38%) 98/177 (54%)	2e-30

- 5 In a BLAST search of public sequence databases, the NOV28a protein was found to have homology to the proteins shown in the BLASTP data in Table 28E.

Table 28E. Public BLASTP Results for NOV28a				
Protein Accession Number	Protein/Organism/Length	NOV28a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P04003	C4b-binding protein alpha chain precursor (C4bp) (Proline-rich protein) (PRP) - <i>Homo sapiens</i> (Human), 597 aa.	1..218 5..224	202/220 (91%) 205/220 (92%)	e-120
Q28065	C4b-binding protein alpha chain precursor (C4bp) - <i>Bos taurus</i> (Bovine), 610 aa.	1..211 5..217	127/214 (59%) 154/214 (71%)	5e-71
S53711	C4BP alpha chain precursor - rabbit, 597 aa.	1..211 5..217	124/214 (57%) 152/214 (70%)	5e-68
P08607	C4b-binding protein precursor (C4bp) - <i>Mus musculus</i> (Mouse), 469 aa.	5..200 17..210	107/196 (54%) 131/196 (66%)	5e-59
Q91X48	Complement component 4 binding protein - <i>Mus musculus</i> (Mouse), 469 aa.	5..200 17..210	107/196 (54%) 130/196 (65%)	8e-59

PFam analysis predicts that the NOV28a protein contains the domains shown in Table 28F.

Table 28F. Domain Analysis of NOV28a			
Pfam Domain	NOV28a Match Region	Identities/ Similarities for the Matched Region	Expect Value
sushi	46..104	16/68 (24%) 42/68 (62%)	1.3e-10
sushi	109..166	20/64 (31%) 47/64 (73%)	6.2e-14
sushi	171..216	20/64 (31%) 38/64 (59%)	0.012

Example 29.

- 5 The NOV29 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 29A.

Table 29A. NOV29 Sequence Analysis

	SEQ ID NO: 99	1356 bp	
NOV29a, CG171681-01 DNA Sequence	CTGCGCTGCCGAGGCGAGCTAAGCGCCCGCTCGCCATGGGGAGCCCCGCACATCGGCCCG CGCTGCTGCTGCTGCTGCGCCTCTGCTGCTGCTGCTGCTGCTGCGCGTCCCGCCAGCC GCAGCTTCCCAGATATGGAACCTCCTAGAATCAAGTGCCCAAGTGTGAAGGAACGCATTG CAGAACCCAACAACTGACAGTCCGGGTGTCTGGGAGACACCCGAAGGAAGAGACACAG CAGATGGAATTCTTACTGATGTCATTCTAAAAGGCCTCCCCCAGGCTCCAACCTTCCAG AAGGAGACCACAAGATCCAGTACACAGTCTATGACAGAGCTGAGAATAAGGGCACTTGCA AATTTCGAGTTAAAGTAAGAGTCAAACGCTGTGGCAAACCTCAATGCCCCAGAGAATGGTT ACATGAAGTGCTCCAGCGACGGTGATAATTATGGAGCCACCTGTGAGTTCTCCTGCATCG GCGGCTATGAGCTCCAGGGTAGCCCTGCCCGAGTATGTCAATCCAACCTGGCTTGGTCTG GCACGGAGCCCACCTGTGCAGCCATGAACGTCAATGTGGGTGTGAGAACGGCAGCTGCAC TTCTGGATCAGTTTATGAGAAAAGGAGACTCCTCATTGTGCCACACCCACAGCCCGAA ACCTCCTTTACCGGCTCCAGCTAGGAATGCTGCAGCAAGCACAGTGTGGCCTTGATCTTC GACACATCACCGTGGTGGAGCTGGTGGGTGTGTTCCCGACTCTCATTGGCAGGATAGGAG CAAAGATTATGCCTCCAGCCCTAGCGCTGCAGCTCAGGCTGTGCTGCGAATCCCCTCT ACTCCTTCAGTATGGTGTAGTGGATAAGCATGGCATGGACAAAGAGCGCTATGTCTCCC TGGTGATGCTGTGGCCCTGTTCAACCTGATTGACACTTTTCCCTTGAGAAAAGAAGAGA TGGTCTTACAAGCCGAAATGAGCCAGACCTGTAACACCTGACATGATGGTCTCTCTTG GCAATCCTCTTATTGTCTACATAGTGACATGCACACGGGAAAGCCTTAAAAATATCCT TGATGTACAGATTTATTTGTAATTTTAAAAGTCTATTTTATTATGAGCTTTCTTTGCAC TTAATAATTAGCATGCTGCTTTTGTACTTGGAAAGTGTTCAAAAAATTATATGACCATA TTTACTCTTTTAACCTTTCTTTACTCCATCATGGCTGGTTGATTGTAGAGAAATTAGA ACCCATAACCATAACAGGCTATCAACATGTTATTCAATGTGACACCTAACTCTTTCTA TTTTGTTTTTTAAGTAAGACTTTTATTAATAAAACG		
	ORF Start: ATG at 36		ORF Stop: TGA at 999
	SEQ ID NO: 100	321 aa	MW at 35636.4kD
NOV29a, CG171681-01 Protein Sequence	MGSPAHRPALLLLLPPLLLLLLRLVPPSRSPDMEPPRIKCPSVKERIAEPNKLTVRVSW ETPEGRDADGILTDVILKGLPPGSNPFEGDHKIQYTVYDRAENKGTCKFRVKVRKRCG KLNAPENGYMKSSDGDNDYGATCEFSICGGYELQGS PARVCQSNLWSGTEPTCAAMNVN VGVRTAAALLDQFYEKRRLLIVSTPTARNLLYRLQLGMLQQAQCGLDLRHITVVELVGVF PTLIGRIGAKIMPPALALQLRLLLRIPLYSFSMVLVDKHGMDKERYVSLVMPVALFNLID TFPLRKEEMVLQAEMSQTCNT		
	SEQ ID NO: 101	1795 bp	
NOV29b, CG171681-03 DNA Sequence	CTTGGTCTCTTCGGTCTCCTGCCGCCCGGGGAAGCGCGTGCCTGCCGAGGCGAGCTA AGCGCCCGCTCGCCATGGGGAGCCCCGCACATCGCCCGCGTGTGCTGCTGCTGCCGC CTCTGCTGCTGCTGCTGCTGCGCGTCCCGCCAGCCGAGCTTCCCAGATACCCCGTGGT GCTCCCCCATCAAGGTGAAGTATGGGGATGTGTACTGCAGGGCCCTCAAGGAGGATACT ACAAACAGCCCTGGGAACCAAGGTGCACATTCTGCTGCCAGAAGGGCTACGAGCTGCATG GCTCTTCCCTACTGATCTGCCAGTCAAACAAACGATGGTCTGACAAGGTATCTGCAAC AAAAGCGATGTCTACCTTGCCATGCCAGCAAATGGAGGGTTAAGTGTGTAGATGGTG CCTACTTTAACTCCCGGTGTGAGTATTATTGTTACCAGGATACACGTTGAAAGGGGAGC GGACCGTCACATGTATGGACAACAAGGCCTGGAGCGGCCGCGCAGCCTCCTGTGTGGATA TGGAACTCCTAGAATCAAGTGCCCAAGTGTGAAGGAACGCATTGCAGAACCAACAAAC TGACAGTCCGGGTGTCTGGGAGACACCCGAAGGAAGAGACACAGCAGATGGAATTCTTA CTGATGTCATTCTAAAAGGCCTCCCCCAGGCTCCAACCTTCCAGAAGGAGACCACAAGA TCCAGTACACAGTCTATGACAGAGCTGAGAATAAGGGCACTTGCAAATTTGAGTTAAAG TAAGAGTCAAACGCTGTGGCAAACCTCAATGCCCCAGAGAATGGTTACATGAAGTGTCCA GCGACGGTGATAATTATGGAGCCACCTGTGAGTTCTCCTGCATCGGCGGCTATGAGCTCC AGGGTAGCCCTGCCCGAGTATGTCAATCCAACCTGGCTTGGTCTGGCACGGAGCCACCT GTGCAGCCATGAACGTCAATGTGGGTGTGAGAACGGCAGCTGCACTTCTGGATCAGTTT ATGAGAAAAGGAGACTCCTCATTGTGTCCACACCCACAGCCGAAACCTCCTTTACCGGC TCCAGCTAGGAATGCTGCAGCAAGCACAGTGTGGCCTTGATCTTCGACACATCACCGTGG TGGAGCTGGTGGGTGTGTTCCCGACTCTCATTGGCAGGATAGGAGCAAAGATTATGCCTC CAGCCCTAGCGCTGCAGCTCAGGCTGTTGCTGCGAATCCCCTCTACTCCTTCAGTATGG TGCTAGTGGATAAGCATGGCATGGACAAAGAGCGCTATGTCTCCTGGTGTGCTGCTGG CCTGTTCACCTGATTGACACTTTCCCTTGAGAAAAGAAGAGATGGTCTCAACAGCCG AAATGAGCCAGACCTGTAACACCTGACATGATGGTTCCTCTCTTGGCAATTCCTCTTCAT		

	TGTCTACATAGTGACATGCACACGGGAAAGCCTTAAAAATATCCTTGATGTACAGATTTT ATTTGTAATTTTAAAAGTCTATTTTATTATGAGCTTTCTTTGCACCTTAAAAATTAGCATG CTGCTTTTTGTACTTGGAAGTGTTTCAAAAAATTATATGACCATATTACTCTTTCTAAC TTTCTTTACTCCATCATGGCTGGTTGATTTGTAGAGAAATTAGAAGCCATAACCATACA CAGGCTATCAACATGTTATTCAATGTGACACCTAACTCTTTCTATTTTGTTTTAAAGT AAGACTTTTATTAATAAAACAAAATGTTTTGGAGCAAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 75		ORF Stop: TGA at 1404
	SEQ ID NO: 102	443 aa	MW at 49267.9kD
NOV29b, CG171681-03 Protein Sequence	MGSPAHRPALLLLPPLLLLLLRVPPSRSPDTPWCSPKVKYGDVYCRAPQGGYYKTAL GTRCDIRCQGYELHGSSLLICQSNKRWSKVICQKRCPTLAMPANGGFKCVDGAYFNS RCEYYCSPGYTLKGERTVTCMDNKAWSGRPASCVDMPEPRKICPSVKERIAEPNKLTVRV SWETPEGRDADGILTDVILKGLPPGSNPFEGDHKIQYTVYDRAENKGTCKFRVKVRVKR CGKLNAPENGYMKSSDGDNYGATCEFSICGGYELQGS PARVCQSNLAWSGTEPTCAAMN NVNVGVRTAAALLDQFYEKRRLLIVSTPTARNLLYRLQLGMLQQAQCGLDLRHITVVELVG VFPTLIGRIGAKIMPPALALQLRLLRIPLYSFSMVLVDKHGMDKERYVSLVMPVALFNL IDTFPLRKEEMVLQAEMSQTCNT		
	SEQ ID NO: 103	1798 bp	
NOV29c, CG171681-02 DNA Sequence	CTGGGTCTCTCGGTCTCCTGCCGCCCGGGAAGCGCGCTGCGCTGCCGAGGCGAGCTA AGCGCCCGCTCGCCATGGGAGCCCCGCACATCGGCCCGCGCTGCTGCTGCTGCCGC CTCTGCTGCTGCTGCTGCTGCTGCGCGTCCGCCAGCCGAGCTTCCAGATACCCCGT GGTGCTCCCCCATCAAGGTGAAGTATGGGATGTGTACTGCAGGGCCCTCAAGGAGGAT ACTACAAAACAGCCCTGGGAACAGGTGCGACATTGCTGCCAGAAGGGCTACGAGCTGC ATGGCTCTTCCCTACTGATCTGCCAGTCAAACAAACGATGGTCTGACAAGGTCATCTGCA AACAAAAGCGATGTCTACCTTGCCATGCCAGCAAATGGAGGGTTTAAAGTGTGATAGT GTGCCTACTTTAACTCCCGGTGTGAGTATTATTGTTTACCAGGATACAGTTGAAAGGGG AGCGGACCGTCACATGTATGGACAACAAGGCTGGAGCGGCCGCCAGCCTCCTGTGTGG ATATGGAACCTCCTAGAATCAAGTGCCCAAGTGTGAAGGAACGCATTGCAGAACCCAACA AACTGACAGTCCGGGTGTCTGGGAGACACCCGAAGGAAGAGACACAGCAGATGGAATTC TTACTGATGTCAATTCTAAAAGGCTCCCCCAGGCTCCAACTTTCCAGAAGGAGACCACA AGATCCAGTACACAGTCTATGACAGAGCTGAGAATAAGGGCACTTGCAAATTTGAGTTA AAGTAAGAGTCAAACGCTGTGGCAAACTCAATGCCCCAGAGAATGGTTACATGAAGTGCT CCAGCGACGGTGATAATTATGGAGCCACCTGTGAGTTCTCCTGCATCGGCGGCTATGAGC TCCAGGGTAGCCCTGCCGAGTATGTCAATCCAACCTGGCTTGGTCTGGCACGGAGCCCA CCTGTGCAGCCATGAACGTCAATGTGGGTGTCAGAACGGCAGCTGCACCTTCTGGATCAGT TTTATGAGAAAAGGAGACTCCTCATTGTGTCCACACCCACAGCCGAAACCTCCTTTACC GGCTCCAGCTAGGAATGTGTCAGCAAGCACAGTGTGGCCTTGATCTTCGACACATCACCG TGGTGGAGCTGGTGGGTGTGTTCCCGACTCTCATTGGCAGGATAGGAGCAAAGATTATGC CTCCAGCCCTAGCGCTGCAGCTCAGGCTGTTGCTGCGAATCCACTTACTCCTTCAGTA TGGTGCTAGTGATAAGCATGGCATGGACAAAGAGCGCTATGTCTCCTGGTGTGCTG TGGCCCTGTTCAACCTGATTGACACTTTTCCCTTGAGAAAAGAAGAGATGGTCTACAAG CCGAAATGAGCCAGACCTGTAACCTGACATGATGGTTCTCTCTTGGCAATTCTCTT CATTTGTCTACATAGTGACATGCACACGGGAAAGCCTTAAAAATATCCTTGATGTACAGAT TTTATTTGTAATTTTAAAAGTCTATTTTATTATGAGCTTTCTTTGCACCTTAAAAATTAGC ATGCTGCTTTTGTACTTGGAAGTGTTTCAAAAAATTATATGACCATATTTACTCTTTCT AACTTTCTTTACTCCATCATGGCTGGTTGATTTTGTAGAGAAATTAGAAGCCATAACCAT ACACAGGCTATCAACATGTTATTCAATGTGACACCTAACTCTTTTCTATTTTGTTTTAA AGTAAGACTTTTATTAATAAAACAAAATGTTTTGGAGCAAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 75		ORF Stop: TGA at 1407
	SEQ ID NO: 104	444 aa	MW at 49381.1kD
NOV29c, CG171681-02 Protein Sequence	MGSPAHRPALLLLPPLLLLLLRVPPSRSPDTPWCSPKVKYGDVYCRAPQGGYYKTA LGTRCDIRCQGYELHGSSLLICQSNKRWSKVICQKRCPTLAMPANGGFKCVDGAYFN SRCEYYCSPGYTLKGERTVTCMDNKAWSGRPASCVDMPEPRKICPSVKERIAEPNKLTVR VSWETPEGRDADGILTDVILKGLPPGSNPFEGDHKIQYTVYDRAENKGTCKFRVKVRVK RCGLNAPENGYMKSSDGDNYGATCEFSICGGYELQGS PARVCQSNLAWSGTEPTCAAM NVNVGVRTAAALLDQFYEKRRLLIVSTPTARNLLYRLQLGMLQQAQCGLDLRHITVVELV GVFPPTLIGRIGAKIMPPALALQLRLLRIPLYSFSMVLVDKHGMDKERYVSLVMPVALFN LIDTFPLRKEEMVLQAEMSQTCNT		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 29B.

Table 29B. Comparison of NOV29a against NOV29b and NOV29c.		
Protein Sequence	NOV29a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV29b	33..321 155..443	273/289 (94%) 273/289 (94%)
NOV29c	33..321 156..444	273/289 (94%) 273/289 (94%)

Two polymorphic variants of NOV29c have been identified and are shown in Table 41K.

- 5 Further analysis of the NOV29a protein yielded the following properties shown in Table 29C.

Table 29C. Protein Sequence Properties NOV29a	
PSort analysis:	0.8200 probability located in outside; 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in lysosome (lumen)
SignalP analysis:	Cleavage site between residues 31 and 32

A search of the NOV29a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 29D.

10

Table 29D. Geneseq Results for NOV29a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV29a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB07747	A human cancer-associated protein-1 (CAP-1) - <i>Homo sapiens</i> , 465 aa. [WO200043508-A2, 27-JUL-2000]	33..319 178..464	148/287 (51%) 205/287 (70%)	7e-89

AAB59009	Breast and ovarian cancer associated antigen protein sequence SEQ ID 717 - <i>Homo sapiens</i> , 431 aa. [WO200055173-A1, 21-SEP-2000]	33..319 144..430	148/287 (51%) 205/287 (70%)	7e-89
ABB72149	Rat protein isolated from skin cells SEQ ID NO: 188 - <i>Rattus</i> sp, 118 aa. [WO200190357-A1, 29-NOV-2001]	88..203 3..118	71/116 (61%) 89/116 (76%)	3e-38
AAB55949	Skin cell protein, SEQ ID NO: 188 - <i>Rattus</i> sp, 118 aa. [WO200069884-A2, 23-NOV-2000]	88..203 3..118	71/116 (61%) 89/116 (76%)	3e-38
AA Y76010	Rat DRS protein homolog, SEQ ID NO:188 - <i>Rattus</i> sp, 118 aa. [WO9955865-A1, 04-NOV-1999]	88..203 3..118	71/116 (61%) 89/116 (76%)	3e-38

In a BLAST search of public sequence databases, the NOV29a protein was found to have homology to the proteins shown in the BLASTP data in Table 29E.

Table 29E. Public BLASTP Results for NOV29a				
Protein Accession Number	Protein/Organism/Length	NOV29a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P78539	Sushi repeat-containing protein SRPX precursor - <i>Homo sapiens</i> (Human), 464 aa.	33..321 176..464	289/289 (100%) 289/289 (100%)	e-168
Q63769	Sushi repeat-containing protein SRPX precursor (DRS protein) (Down-regulated by V-SRC) - <i>Rattus norvegicus</i> (Rat), 464 aa.	33..321 176..464	279/289 (96%) 286/289 (98%)	c-164
Q9R0 m3	Sushi-repeat-containing protein - <i>Mus musculus</i> (Mouse), 464 aa.	33..320 176..463	276/288 (95%) 285/288 (98%)	e-163
Q9R0 m2	Sushi-repeat-containing protein - <i>Mus musculus</i> (Mouse), 380 aa.	33..320 92..379	276/288 (95%) 285/288 (98%)	c-163

AAM73690	Sushi-repeat containing protein - <i>Mus musculus</i> (Mouse), 410 aa (fragment).	33..319 123..409	152/287 (52%) 203/287 (69%)	2e-89
----------	---	---------------------	--------------------------------	-------

PFam analysis predicts that the NOV29a protein contains the domains shown in Table 29F.

Table 29F. Domain Analysis of NOV29a			
Pfam Domain	NOV29a Match Region	Identities/ Similarities for the Matched Region	Expect Value
HYR	33..114	27/86 (31%) 78/86 (91%)	2.2e-34
sushi	119..174	19/64 (30%) 41/64 (64%)	2.7e-09

Example 30.

The NOV30 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 30A.

Table 30A. NOV30 Sequence Analysis			
	SEQ ID NO: 105	1499 bp	
NOV30a, CG51117-01 DNA Sequence	ACGCGTGTAGGTGGCCAGGCAAAATAGTGTCATCGATTGGCCTATGTCGTTATGGTGGGA GGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTACGCTTTCTACG TCTTAAGGCAGAGAATAGCCAGGATAAGGTGCCAGCTCAAAGCTGTGTGCCAACCCAGAT GCAAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTCTCCTGGTTATGCTGGAA AAACCTGGTATTCAGTTTAAATGAGTGTGGCCTGAAGCCCGGCCCTGTAAGCAGAGG TGCATGAACACTTACGGCAGCTACAGTGCTACTGTCTCAACGGATATATGCTCATGCCG GATGGTTCTGCTCAAGTGCCTGACCTGCTCCATGGCAAAGTGTGAGTATGGCTGTGAT GTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCTGGCCTGCAGCTGGCTCCTGAT GGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATTT AGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTCAAAAGGCTTCGATCTC ATGTATATTGGAGGCAATATCAATGTCTGACATAGACGAATGCTCACTTGGTCACTAT CAGTGCAGCAGCTTTGCTCGATGTTATAACGTACGTGGGTCTACAAGTGCAAAATGTAAG GAAGGATACAGGGTGATGGACTGACTTGTGTGTATATCCAAAAGTTATGATTGAACCT TCAGGTCCAATTCATGTACCAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAAT AATAATTGGATTCCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCT CCTATCATTACCAACAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAGCCAACA CCAATTCTACTCCACCACCACCACCACCCCTGCCAACAGAGCTCAGAACACCTTACCA CCTACAACCCAGAAAAGGCCAACCCGGAAGTACAACTATAGCAGCAGCTGCCAGTACA CCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCTCAGAAAACCCAGAGGA GATGTGTTTCAATCCACGGCAACCTTCAAATGACTTGTGTTGAAATATTTGAAATAGAAAG GGAGTCAGTGCAGACGATGAAGCAAAGGATGATCCAGGTGTTCTGGTACACAGTTGTAAT TTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAATGACTTGCAGTGGGAACCA ATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTGCGCAGCCAAAGCCCCAGGGGA AAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCTCATGCATTACAGGGACCTGTGCTG TCATTACGGCACAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAA		
	ORF Start: at 148		ORF Stop: at 1498
	SEQ ID NO: 106	450 aa	MW at 48855.5kD

NOV30a, CG51117-01 Protein Sequence	GASSKLCANHDANMVNVSGQTSASVILVMLEKPGIQVLNECGLKPRPCKHRCMNTYGSYK CYCLNGYMLMPDGSCSSALTCSMANQYGCDDVVKQIRQCPSGLQLAPDGRCTCDVDE CATGRASCPFRQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDEC SLGQYQCSSFARCY NVRGSYKCKCKEGYQDGLTCVYIPKVMIEPSGPIHVPKNGTILKGD TGNNNWI PDVGS TWWPPKTPYIPPIITNRPTSKPTTRPTPKPTPIPTPPPPPLPTELRPLPTTPTTPTT GLTTIAPAASTPPGGITVDNRVQTD PQKPRGDVFI PRQPSNDLFEI FEIERGVSADDEAK DDPGVLVHSCNFDHGLCGWIREKDN DLHWEPIRDPAGGQYLT VSAAKAPGGKAARLVLP L GRLMHSGDLCLSFRHKVTGLHSGTLQVFVR		
	SEQ ID NO: 107	1638 bp	
NOV30b, CG51117-05 DNA Sequence	GAGTTCGACGGGAGGTGGCCAGGCAAATAGTGT CATCGATTGGCCTATGTCGTTATGGT GGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGT CAGCCTGTG TGCCAACCACGATGCAAAATGCTGTAATGTATCGGGCCAAACAAGTGCAAGTGT CATCCT GGTTATGCTGGAAAAACCTGTATTCAAGTTTAAATGAGTGTGGCCTGAAGCCCCGGCCC TGTAAGCACAGGTGCATGAACACTTACGGCAGCTACAAGTGCTACTGTCTCAACGGATAT ATGCTCATGCCGGATGGTTCCTGCTCAAGTGCCCTGACCTGCTCCATGGCAA ACTGTCAG TATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCTTGGCCTGCAG CTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCC TGCCCTAGATTTAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGT CATAAA GGCTTCGATCTCATGTATATTTGGAGGCAAATATCAATGT CATGACATAGACGAATGCTCA CTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAACGTACGTGGGTCTTACAAG TGCAATGTAAAGAAGGATACAGGGTGATGGACTGACTTGTGTGTATATCCCAAAAGTT ATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGGGAAATGGTACCATTTTAAAGGT GACACAGGAAATAATAATTGGATTCTGTATGTTGGAAGTACTTGGTGGCCTCCGAAGACA CCATATATTCTCTCTATCATTTACCAACAGGCCTACTTCTAAGCCAACAAGACCTACA CCAAAGCCAACACCAATTCCTACTCCACCACCACCACCACCCTGCCAACAGAGCTCAGA ACACCTCTACCACCTACAACCCAGAAAGGCCAACACCGGACTGACAACTATAGCACC GCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCCTCAG AAACCCAGAGGAGATGTGTTCAATCCACGGCAACCTTCAAATGACTTGTTTGAAATATTT GAAATAGAAAGAGGAGTCAGTGCAGACGATGAAGCAAAGGATGATCCAGGTGTTCTGGTA CACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAATGACTTG CACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCCGCAGCCAAA GCCCCAGGGGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCCTTATGCATT CAGGG GACCTGTGCCTGTCAATCAGGCACAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTG TTTGTGAGAAAACACGGTGCCACGGAGCAGCCCTGTGGGGAAGAAATGGTGGCCATGGC TGGAGGCAAAACACAGATCACCTTGCAGGGGCTGACATCAAGAGCGTCGTCTTCAAAGGT GAAAAAAGGCGTGGTCACACTGGGGAGATTGGATTAGATGATGTGAGCTTGAAAAAAGGC CACTGCTCTGAAGAACGC		
	ORF Start: at 1	ORF Stop: end of sequence	
	SEQ ID NO: 108	546 aa	MW at 59854.9kD
NOV30b, CG51117-05 Protein Sequence	EFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCP RCKHCEIGPNKCKCHP GYAGKTCIQVLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALTCSMANQ YGCDDVVKQIRQCPSGLQLAPDGRCTCDVDECATGRASCPFRQCVNTFGSYICKCHK GFDLMYIGGKYQCHDIDEC SLGQYQCSSFARCYNVRGSYKCKCKEGYQDGLTCVYIPKV MIEPSGPIHVPKNGTILKGD TGNNNWI PDVGS TWWPPKTPYIPPIITNRPTSKPTTRPT PKPTPIPTPPPPPLPTELRPLPTTPTTPTTGLTTIAPAASTPPGGITVDNRVQTD PQ KPRGDVFI PRQPSNDLFEI FEIERGVSADDEAKDDPGVLVHSCNFDHGLCGWIREKDN DL HWEPIRDPAGGQYLT VSAAKAPGGKAARLVLP L GRLMHSGDLCLSFRHKVTGLHSGTLQV FVRKHGAHGAALWGRNGGHWRQTQITLRGADIKSVVFKGEKRRGHTGEIGLDDVSLKKG HCSEER		
	SEQ ID NO: 109	2245 bp	
NOV30c, CG51117-06 DNA Sequence	GGACACTGACATGGACTGAAGGAGTAGAAAAGAAGGGAGCGGGAGGGGGCTCCGGGCGCC GCGCAGCAGACTGCTCCGGCCGCGCCTCGCGCTGTCTCCGGGAGCGGCAGAGTA GCCCGGGCGGCGAGGGCTGGGGGTTCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCG CCCACCACCCCAACCTGTCTCTCGCGCGCCACTGCGCTGCGCCCCAGGACCCGCTGCCCA ACATGGATTTTCTCTGGCGCTGGTGCTGGTATCTCTCGCTCTACCTGACGGCGGCGCCG AGTTCGACGGGAGTAGGTGGCCAGGCAAATAGTGT CATCGATTGGCCTATGTCGTTATG GTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGACAGTGT CAGCCTT		

	<p>TCTACGTCTTAAGGCAGAGAATAGCCAGGATAAGGTGCCAGCTCAAAGCTGTGTGCCAAC CACGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTATCCTGGTTATG CTGGAAAAACCTGTAATCAAGACGAGCACATCCCAGCTCCTCTTGACCAAGGCAGTGAAC AGCCTCTTTTCCAACCCCTGGATCACCAGCCACAAGTTTGCCTTCAAGGGATCTAAATG AGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTACGGCAGCTACA AGTGCTACTGTCTCAACGGATATATGCTCATGCCGATGGTTCCTGCTCAAGTGCCCTGA CCTGCTCCATGGCAAACCTGTCTAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCC AGTGCCCATCCCTGGCCTGCAGCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATG AATGTGCTACAGGAAGAGCCTCCTGCCCTAGATTTAGGCAATGTGTCAACACTTTTGGGA GCTACATCTGCAAGTGTATAAAGGCTTCGATCTCATGTATATTGGAGGCAAATATCAAT GTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTT ATAACATACGTGGGTCTTACAAGTGCAAATGTAAGAAGGATACCAGGTGATGGACTGA CTTGTTGTATATCCCAAAAGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGG GAAATGGTACCATTTTAAAGGGTGACACAGGAAATAATAATTGGATTCTGATGTTGGAA GTACTTGGTGGCCTCCGAAGACACCATATATCCTCCTATCATTACCAACAGGCCTACTT CTAAGCCAACAACAAGACCTACACCAAAGCCAACCAATTCCTACTCCACCACCCAC CACCCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCAGAAAGGCCAACCA CCGGACTGACAACTATAGCACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACA ACAGGGTACAGACAGACCTCAGAAACCCAGAGGAGATGTGTTCAATCCACGGCAACCTT CAAATGACTTGTGTTGAAATATTTGAAATAGAAAGAGGAGTCAGTGCAGACGATGAAGCAA AGGATGATCCAGGTGTTCTGGTACACAGTTGTAATTTGACCATGGACTTTGTGGATGGA TCAGGGAGAAAGACAATGACTTGCAGTGGGAACCAATCAGGGACCCAGCAGGTGGACAAT ATCTGACAGTGTGGCAGCCAAAGCCCCAGGGGAAAAGCTGCACGCTTGGTGTACCTC TCGGCCGCTTATGCATTACAGGGACCTGTGCCTGTCAATCAGGCACAAGGTGACGGGGC TGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAACACGGTGCCACGGAGCAGCCCTGT GGGGAAGAAATGGTGGCCATGGCTGGAGGCAACACAGATCACCTTGGAGGGGCTGACA TCAAGAGCGTCGTCTTCAAAGGTGAAAAAAGCGTGGTCACACTGGGGAGATTGGATTAG ATGATGTGAGCTTGA AAAAAGGCCACTGCTCTGAAGAACGCTAACTCCAGAACTAAC AATGAATCCTATGTTGCTCTATCCTCTTTTCCAATTCTCATCTTCTCTCTCTCTCC CTTTATCAGGCCTAGGAGAAGAGTGGGTGAGTGGGTGAGAAGGAAGTCTATTTGGTGAC CCAGGTTCTTCTGGCCTGCTTTTGT</p>		
	ORF Start: ATG at 243		ORF Stop: TAA at 2082
	SEQ ID NO: 110	613 aa	MW at 67416.5kD
NOV30c, CG51117-06 Protein Sequence	<p>MDFLALVLVSSLYLQAAEFDGSRWPRQIVSSIGLCRYGGRIDCCWGWARSWSGQCQPF YVLRQRIARIRCQLKAVCQPRCKHGEICIPNCKKCHPGYAGKTCNQDEHIPAPLDQGSEQ PLFQPLDQHQATSLPSRDLNECGLKPRPCKHRCMNITYGSYKCYCLNGYMLMPDGSCSALT CSMANCQYGCDDVVKQIRQCPSPLQLAPDGRCTVDVDECATGRASCPFRQCNTFGS YICKCHKGFDLMIYIGGKYQCHDIDECSLGQYQCSSFARCYNIRGSYKCKKEGYQDGLT CVYIPKVMIEPSGPIHVPKNGTILKGDGTNNNWIIPDVGSTWPPKTPYIPPIITNRPTS KPTTRPTPKPTPIPTPPPPPLPLTELRTPLPPTTPTTGLTTIAPAASTPPGGITVDN RVQTDPPQKPRGDVFIIPRQPSNDLFEIFEIERGVSADDEAKDDPGVLVHSCNFDHLCGWI REKDNLDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPLGRLMHSGDLCLSFHVKVTGL HSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIKSVVFKGEKRRGHTGEIGLD DVSLKKGHCSEER</p>		
	SEQ ID NO: 111	2194 bp	
NOV30d, CG51117-07 DNA Sequence	<p>GGCACTGACATGGAAGGAGTAGAAAAGAAGGGAGCGGGAGGGGGCTCCGGGCGCC GCGCAGCAGACCTGCTCCGGCCGCGCGCTCGCCGCTGTCTCCGGGAGCGGCAGAGTA GCCCGGGCGGGAGGGCTGGGGTTCCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCG CCCACCAACCACTGTTCTCGCGGCCACTGCGCTGCGCCCGGAGCCGCTGCCCA ACATGGATTTTCTCTGGCGCTGGTGTGCTGCTGCTCTACCTGCAGCGCGCCGCGG AGTTCGACGGGAGTAGGTGGCCAGGCAATAGTGTATCGATTGGCCTATGTCGTTATG GTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGACAGTGTGAGCCTG TGTGCCAACCACGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTATC CTGGTTATGCTGGAAAAACCTGTAATCAAGACGAGCACATCCCAGCTCCTCTTGACCAAG GCAGTGAACAGCCTCTTTTCCAACCCCTGGATCACCAGCCACAAGTTTGCCTTCAAGGG ATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTACG GCAGCTACAAGTGTACTGTCTCAACGGATATATGCTCATGCCGATGGTTCCTGTCAA GTGCCCTGACCTGCTCCATGGCAAACCTGCAGTATGGCTGTGATGTTGTTAAAGGACAAA</p>		

	TACGGTGCCAGTGCCCATCCCTGGCCTGCAGCTGGCTCCTGATGGGAGGACCTGTGTAG ATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATTTAGGCAATGTGTCAACA CTTTTGGGAGCTACATCTGCAAGTGTCTATAAAGGCTTCGATCTCATGTATATTGGAGGCA AATATCAATGTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTG CTCGATGTTATAACATACGTGGGTCTACAAGTGCAAATGTAAAGAAGGATACCAAGGTG ATGGACTGACTTGTGTATATCCCAAAAGTTATGATTGAACCTTCAGTCCAATTTCATG TACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAATAATAATTGGATTCTCTG ATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCTCTATCATTACCAACA GGCCTACTTCTAAGCCAACAACAAGACCTACACCAAGCCAACACCAATTCTACTCCAC CACCACCACCACCCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCAGAGAA GGCCAACCACCGGACTGACAACTATAGCACCAGCTGCCAGTACACCTCCAGGAGGATTA CAGTTGACAACAGGGTACAGACAGACCCTCAGAAACCCAGAGGAGATGTGTTTATTCCAC GGCAACCTTCAAATGACTTGTGTTGAAATATTTGAAATAGAAAGAGGAGTCACTGCAGACG ATGAAGCAAAGGATGATCCAGGTGTTCTGGTACACAGTTGTAATTTTGACCATGGACTTT GTGGATGGATCAGGGAGAAAGACAATGACTTGCACTGGGAACCAATCAGGGACCCAGCAG GTGGACAATATCTGACAGTGTGGCAGCCAAAGCCCCAGGGGAAAAGCTGCACGCTTGG TGCTACCTCTCGGCCGCCTTATGCATTAGGGGACCTGTGCCTGTCTATTAGGCACAAGG TGACGGGGCTGCACTCTGGCACACTCCAGGTGTTGTGAGAAAACACGCTGCCACGGAG CAGCCCTGTGGGGAAGAAATGGTGGCCATGGCTGGAGGCAACACAGATCACTTGCAG GGGCTGACATCAAGAGCGTCGTCTTCAAAGGTGAAAAAGGCGTGGTCACTGGGGAGA TTGGATTAGATGATGTGAGCTTGAAAAAGGCCACTGCTCTGAAGAACCGCTAACAACTCC AGAATAACAATGAACCTCTATGTTGCTCTATCTCTTTTTTCCAATTCTCATCTTCTCTC CTCTTCTCCCTTTTATCAGGCCTAGGAGAAGAGTGGGTCACTGGGTGAGAAGGAAGTCTA TTTGGTGACCCAGGTTCTTCTGGCCTGCTTTTGT		
	ORF Start: ATG at 243	-	ORF Stop: TAA at 2031
	SEQ ID NO: 112	596 aa	MW at 65299.9kD
NOV30d, CG51117-07 Protein Sequence	MDFLALVLVSSLYLQAAEFDGSRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPV CQPRCKHGEICIPNKCKCHPGYAGKTCNQDEHI PAPLDQGSQPLFQPLDHQATSLPSRD LNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGCSSALTCSMANCQYGCDDVVKQI RCQCPSPGLQLAPDGRCTCDVDECATGRASCPRFRQCVNTFGSYICKCHKGFDLMYIGGK YQCHDIDECSLGQYQCSSFARCYNIRGSYKCKCKEGYQGDGLTCVYIPKVMIEPSGPIHV PKNGNTILKGD TGNNWIPDVGSTWVPPKTPYIPPIITNRPTSKPTTRPTKPTPIPTPP PPPPLTELRTPLPPTTPTTGLTTIAPAASTPPGGITVDNRVQTDPOKPRGDVFI PR QPSNDLFEIFEIERGVSADEAKDDPGVLVHSCNFDHGLCGWIREKDNLDHWEPIRDPAG QYLTVSAAKAPGGKAARLVPLGRMLHSGDLCLFRHKVTGLHSGTGLQVFRKHGAHA ALWGRNGGHGWRQTQITLRGADIKSVVFKGEKRRGHTGEIGLDDVSLKKGHCSEER		
	SEQ ID NO: 113	2112 bp	
NOV30e, CG51117-03 DNA Sequence	GGGAGGGGGCTCCGGGCGCGCAGCAGACCTGCTCCGGCGCGCGCTCGCGCTGTGTC CTCCGGGAGCGCAGCAGTAGCCCGGGCGCGAGGGCTGGGGTTCTCTGAGACTCTCAG AGGGGCGCTCCCATCGGCGCCACCAACCACTGTTCTCGCGCGCACTGCGCTGC GCCCAGGACCCGCTGCCAATGATTTTCTCTGGCGCTGGTGCTGGTATCCTCGCT CTACCTGCAGGCGCGCGCGAGTTCGACGGGAGGTGGCCAGGCAATAGTGTATCGAT TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTG GGACAGTGTGAGCCTTTCTACGTCTTAAGGCAGAGAATAGCCAGGATAAGGTGCCAGCT CAAAGCTGTGTGCCAACACGATGCAACATGGTGAATGTATCGGGCCAAACAAGTGCAA GTGTCATCCTGTTATGCTGGAAAAACCTGTATTCAAAGTTTAAATAGTGTGGCCTGAA GCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTACGGCAGCTACAAGTGTACTGTCT CAACGGATATATGCTCATGCCGATGGTTCTTGCTCAAGTGCCCTGACCTGCTCCATGCT AACTGTGAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCATCCCC TGGCCTGCAGCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGG AAGAGCCTCCTGCCCTAGATTTAGGCAATGTGTCAACACTTTGGGAGCTACATCTGCAA GTGTCATAAAGGCTTCGATCTCATGTATATTGGAGGCAATATCAATGTCATGACATAGA CGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAACGTACGTGG GTCTTACAAGTGCAAAATGTAAAGAAGGATACAGGGTATGGACTGACTTGTGTGATAT CCCAAAAGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAGGGAAATGGTACCAT TTTAAAGGGTGACACAGGAAATAATAATTGGATTCTGATGTTGGAAGTACTTGGTGGCC TCCGAAGACACCATATATCTCTCTATCATTACCAACAGGCCTACTTCTAAGCCAACAAC AAGACCTACACCAAGCCAACACCAATTCTACTCCACCACCACCACCACCCTGCCAAC		

	AGAGCTCAGAACACCTCTACCACCTACAACCCAGAAAGGCCAACCCCGGACTGACAAC TATAGCACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGAC AGACCCTCAGAAACCCAGAGGAGATGTGTTTCATCCACGGCAACCTTCAAATGACTTGTT TGAAATATTTGAAATAGAAAGAGGAGTCAGTGCAGACGATGAAGCAAAGGATGATCCAGG TGTTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGA CAATGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTC GGCAGCCAAAGCCCCAGGGGAAAAGCTGCACGCTTGGTGTCTACCTCTCGGCCGCTTAT GCATTCAGGGGACCTGTGCCTGTCTATTAGGCACAAGGTGACGGGGTGCCTCTGGCAC ACTCCAGGTGTTTGTGAGAAAACACGGTGCCACGGAGCAGCCCTGTGGGAAGAAATGG TGGCCATGGCTGGAGGCAACACAGATCACCTTGCAGGGGCTGACATCAAGAGCGTCGT CTTCAAAGGTGAAAAAAGCGTGGTCACTGGGGAGATTGGATTAGATGATGTGAGCTT GAAAAAAGGCCACTGCTCTGAAGAACGCTAACAACTCCAGAACTAACATGAATCCTAT GTTGCTCTATCCTCTTTTCCAATTCTCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT TAGGAGAAGAGTGGGTGAGTGGGTGAGTGGGTGAGTGGGTGAGTGGGTGAGTGGGTGAGTGGGT GCCTGCTTTTGT		
	ORF Start: ATG at 203		ORF Stop: TAA at 1949
	SEQ ID NO: 114	582 aa	MW at 63991.9kD
NOV30e, CG51117-03 Protein Sequence	MDFLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY VLRQRIARIRQLKAVCQPRCKHGEICGPNKCKCHPGYAGKTCIQVLNECGLKPRPCKHR CMNTYGSYKCYCLNGYMLMPDGSSEALTCMANCQYGCDDVVKQIRQCPSGLQLAPD GRTCDVDDECATGRASCPFRQCVNTFGSYICKCHKGFDMYIGGKYQCHDIDECSLGQY QCSSFARCYNVRGSYKCKCKEGYQGDGLTCVYIPKVMIEPSGPIHVPKNGTILKGDGTGN NNWIPDVGSTWVPPKTPYIPPIITNRPTSKPTTRPTPKPTPIPTPPPPPLPTLRLTLP PTTPTPTTGLTTIAPAATPPGGITVDNRVQTDQPKPRGDVFIQPSNDLFEIFEIER GVSADDEAKDDPGVLVHSCNFDHGLCGWIREKNDLHWEPIRDPAGGQYLTVSAKAPGG KAARLVLPGLRLMHSGDLCLSFHVKVTGLHSGTLQVFVRKHGAHGAALWGRNGGGHWRQT QITLRGADIKSVVFKGEKRRGHTGEIGLDDVSLKKGHCSEER		
	SEQ ID NO: 115	691 bp	
NOV30f, CG51117-02 DNA Sequence	GGGAGGGGGCTCCGGGCGCGCGCAGCAGACCTGCTCCGGCGCGCGCTCGCCGCTGTC CTCCGGGAGCGGCAGCAGTAGCCCGGGCGCGAGGGCTGGGGTTCTCTCGAGACTCTCAG AGGGGCGCTCCCATCGGCGCCACCACCCCAACCTGTTCTCGCGCGCCACTCGCCTGC GCCCCAGGACCCGCTGCCCAACATGGATTTTCTCTGCGCTGGTGTGTTATCCTCGCT CTACCTGCAGGCGCGCGCGAGTACGACGGGAGGTGGCCAGGCAAAAGTGTATCGAT TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTG GGGACAGTGTGAGCCTTTCTACGTCTTAAGGCAGAGAATAGCCAGGATAAGGTGCCAGCT CAAAGCTGTGTGCCAACACGATGCAACATGGTGAATGTATCGGGCAAACAAGTGCAA GTGTCATCCTGGTTATGCTGGA AAAACCTGTAATCAAGCCGTAGGTTTGAAGATGTAT GGTTCAGCCGGGCGCGCTGCTACCTGTAAATCCAGCACTTTGGAAGGCCGAGGCG GGCGGATCACGAGGTGAGGATATCGAGACCATCTGGCTAACACGGTGAAACCCCATCTC TACTAAAAATACAAAAA AAAAAAAAAA		
	ORF Start: ATG at 203		ORF Stop: TAA at 572
	SEQ ID NO: 116	123 aa	MW at 13844.1kD
NOV30f, CG51117-02 Protein Sequence	MDFLLALVLVSSLYLQAAAEYDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPFY VLRQRIARIRQLKAVCQPRCKHGEICGPNKCKCHPGYAGKTCNQAVGFERCMVPAGRRG STL		
	SEQ ID NO: 117	261 bp	
NOV30g, CG51117-04 DNA Sequence	GAGTACGACGGGAGGTGGCCAGGCAAAATAGTGTATCGATTGGCCTATGTCGTTATGGT GGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTGAGCCTGTG TGCCAACCACGATGCAACATGGTGAATGTATCGGGCAAACAAGTGCAAGTGTATCCT GGTATGCTGGA AAAACCTGTAATCAAGCCGTAGGTTTGAAGATGTATGGTTCCAGCC GGGCGCGCTGGCTCTACCTG		
	ORF Start: at 1		ORF Stop: end of sequence
	SEQ ID NO: 118	87 aa	MW at 9707.1kD
NOV30g,	EYDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCPCKHGEICGN		

CG51117-04 Protein Sequence	KCKC HPGYAGKTCNQAVGFERCMVPAGRRGSTL
	SEQ ID NO: 119 1804 bp
NOV30h, CG51117-08 DNA Sequence	CACCGGATCCATGGATTTTCTCCTGGCGCTGGTGCTGGTATCCTCGCTCTACCTGCAGGC GGCCGCCGAGTTCGACGGGAGGTGGCCAGGCAAATAGTGTATCGATTGGCCTATGTCG TTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTCA GCCTGTGTGCCAACACGATGCAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTG TCATCCTGGTTATGCTGGAAAACTGTAATCAAGACGAGCACATCCCAGCTCCTCTTGA CCAAGGCAGTGAACAGCCTCTTTTCCAACCCCTGGATCACCAGCCACAAGTTTGCCTTC AAGGGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACAC TTACGGCAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCTTG CTCAAGTGCCCTGACCTGCTCCATGGCAAACCTGTCTAGTATGGCTGTGATGTTGTTAAAGG ACAAATACGGTGCCAGTGCCCATCCCCTGGCCTGCACCTGGCTCCTGATGGGAGGACCTG TGTAAGTGTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATTTAGGCAATGTGT CAACACTTTTGGGAGCTACATCTGCAAGTGTCTATAAAGGCTTCGATCTCATGTATATTGG AGGCAAATATCAATGTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAG CTTTGCTCGATGTTATAACGTACGTGGGTCTACAAGTGCAAATGTAAAGAAGGATACCA GGGTGATGGACTGACTTGTGTATATCCCAAAGTTATGATTGAACCTTCAGGTCCAAT TCATGTACCAAAGGAAATGGTACCATTTTAAAGGGTGACACAGGAAATAAATTTGGAT TCCTGATGTTGGAAGTACTTGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTAC CAACAGGCCTACTTCTAAGCCAACAACAGACCTACACCAAAGCCAACCAATTCCTAC TCCACCACCACCACCACCCCTGCCAACAGAGCTCAGAACACCTTACCACCTACAACCCC AGAAAGGCCAACACCAGGACTGACAACTATAGCACCAGCTGCCAGTACACCTCCAGGAGG GATTACAGTTGACAACAGGGTACAGACAGACCCCTCAGAAACCAGAGGAGATGTGTTTAT TCCACGGCAACCTTCAAATGACTTGTGTTGAAATATTGAAATAGAAAGAGGAGTCAGTGC AGACGATGAAGCAAAGGATGATCCAGGTGTTCTGGTACACAGTTGTAATTTTGACCATGG ACTTTGTGGATGGATCAGGGAGAAAGACAATGACTTGCCTGGGAACCAATCAGGGACCC AGCAGGTGGACAATATCTGACAGTGTCCGACAGCCAAAGCCCCAGGGGAAAGCTGCACG CTTGGTGCTACCTCTCGGCCGCTCATGCATTACAGGGACCTGTGCCTGTCTTACAGGCA CAAGGTGACGGGGTGCCTCTGGCACACTCCAGGTGTTTGTGAGAAAACAGCGTGCCCA CGGAGCAGCCCTGTGGGGAAGAAATGGTGGCCATGGCTGGAGGCAAAACACAGATCACCTT GCGAGGGGCTGACATCAAGAGCGTCTGCTTCAAAGGTGAAAAAAGCGTGGTCACTGCG GGAGATTGGATTAGATGATGTGAGCTTGAAGAACCGCTGCTCTGAAGAACCGCTCGA CGGC
	ORF Start: ATG at 11 ORF Stop: at 1796
	SEQ ID NO: 120 595 aa MW at 65207.8kD
NOV30h, CG51117-08 Protein Sequence	MDFLALVLVSSLYLQAAEFDRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPV QPRCKHGEICIGPNKCKCHPGYAGKTCNQDEHIPAPLDQGGSEQLFQPLDQATSLPSRDL NECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALTCSMANQYQCDVVKQIR CQCPSPGLHLAPDGRTCVDVDECATGRASCPRFRQCVNTFGSYICKCHKGFDLMYIGGKY QCHDIDECSLGQYQCSSFARCYNVRGSYKCKKEGYQGDGLTCVYIPKVMIEPSGPIHVP KNGTILKGDGTGNWNWIPDVGSTWWPPKTPYIPPIITNRPTSKPTTRPTPTPIPTPPP PPPLPTELRTPLPPTTPTTGLTTIAPAASTPPGGITVDNRVQTDQPKPRGDVFI PSNDLFEIFEIERGVSADDEAKDDPGVLVHSCNFDHGLCGWIREKNDLHWEPIRDPAGG QYLTVSAAKAPGGKAARLVLPGLRLMHSGLCLSPFRHKVTGLHSGTLQVFRKHGAHGA LWGRNGGHGWRQTQITLRGADIKSVVFKGEKRRGHTGEIGLDDVSLKKGHCSEER
	SEQ ID NO: 121 1858 bp
NOV30i, CG51117-09 DNA Sequence	CACCGGATCCATGGATTTTCTCCTGGCGCTGGTGCTGGTATCCTCGCTCTACCTGCAGGC GGCCGCCGAGTTCGACGGGAGTAGGTGGCCAGGCAAATAGTGTATCGATTGGCCTATG TCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTG TCAGCCTTTCTACGTCTTAAGGCAGAGAATAGCCAGGATAAGGTGCCAGCTCAAAGCTGT GTGCCAACACGATGCAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTATCC TGTTATGCTGGAAAACTGTAATCAAGACGAGCACATCCAGCTCCTCTTGACCAAGG CAGTGAACAGCCTCTTTTCCAACCCCTGGATCACCAGCCACAAGTTTGCCTTCAAGGGA TCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTACGG CAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCCTGTCTAAG TGCCCTGACCTGCTCCATGGCAAACCTGTCTAGTATGGCTGTGATGTTGTTAAAGGACAAAT

	ACGGTGCCAGTGCCCATCCCTGGCCTGCAGCTGGCTCCTGATGGGAGGACCTGTGTAGA TGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATTTAGGCAATGTGTCAACAC TTTTGGGAGCTACATCTGCAAGTGTCTATAAAGGCTTCGATCTCATGTATATTGGAGGCAA ATATCAATGTCTAGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGC TCGATGTTATAACGTACGTGGGTCTTACAAGTGCAAATGTAAAGAAGGATACCAGGGTGA TGGACTGACTTGTGTGTATATCCCAAAAGTTATGATTGAACCTTCAGGTCCAATTTCATGT ACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAATAATAATTGGATTCTCTGA TGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAACAG GCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCTACTCCACC ACCACCACCACCCCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCAGAAAG GCCAACACCAGGACTGACAACTATAGCACCAGCTGCCAGTACACCTCCAGGAGGGATTAC AGTTGACAACAGGGGTACAGACAGCCCTCAGAAACCCAGAGAGATGTGTTTCATTCCACG GCAACCTTCAAATGACTTGTGTTGAAATATTTGAAATAGAAAGAGGAGTCAGTGCAGACGA TGAAGCAAAGGATGATCCAGGTGTTCTGGTACACAGTTGTAATTTTGACCATGGACTTTG TGGATGGATCAGGGAGAAAGACAATGACTTGCCTGCGGAACCAATCAGGGACCCAGCAGG TGGACAATATCTGACAGTGTGCGCAGCCAAAGCCCCAGGGGAAAAGCTGCACGCTTGGT GCTACCTCTCGGCCCGCTCATGCATTCAAGGGACCTGTGCCTGTCTATTACGGACAAGGT GACGGGCTGCACTCTGCGCACTCCAGGTGTTTGTGAGAAAACACGGTGCCACAGGAGC AGCCCTGTGGGGAAGAAATGGTGGCCATGGCTGGAGGCAAACACAGATCACCTTGCAGAG GGCTGACATCAAGAGCGTCGTCTTCAAAGGTGAAAAAAGGCGTGGTCACACTGGGGAGAT TGGATTAGATGATGTGAGCTTGAAAAAAGGCCACTGCTCTGAAGAACGCGTCGACGGC	
	ORF Start: ATG at 11	ORF Stop: at 1850
	SEQ ID NO: 122	613 aa MW at 67402.4kD
NOV30i, CG51117-09 Protein Sequence	MDFLLALVLVSSLYLQAAAEFDGSRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPF YVLRQRIARIQQLKAVCQPRCKHGEICIGPNKCKCHPGYAGKTCNQDEHIPAPLDQGSSEQ PLFQPLDHQATSLPSRDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALT CSMANCQYGCDEVVKQIRQCPSPLQLAPDGRCTVDVDECATGRASCPRFRQCVENTFGS YICKCHKGFDLMYIGGKYQCHDIDCSLGQYQCSSFARCYNVRGSYKCKCKEGYQGDGLT CVYIPKVMIEPSGPIHVPKNGNTILKGDGTGNWNWIPDVGSTWVWPPTPYIPPIITNRPST KPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDN RVQTDPPQKPRGDVFI PRQPSNDLFEIFEIERGVSADEAKDDPGVLVHSCNFDHGLCGWI REKNDNLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPGLRLMHSGLCLLSFRHKVTGL HSGTLQVFRKHGAHGAALWGRNGHGWRTQITLRGADIKSVVFKGEKRRGHTGEIGLD DVSLLKKGHCSEER	

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 30B.

Table 30B. Comparison of NOV30a against NOV30b through NOV30i.		
Protein Sequence	NOV30a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV30b	32..240 65..273	207/209 (99%) 207/209 (99%)
NOV30c	1..240 98..340	210/244 (86%) 217/244 (88%)
NOV30d	1..240 81..323	210/244 (86%) 217/244 (88%)
NOV30e	32..240 101..309	207/209 (99%) 207/209 (99%)

NOV30f	184..196 88..100	8/13 (61%) 8/13 (61%)
NOV30g	167..196 33..64	14/32 (43%) 15/32 (46%)
NOV30h	1..240 80..322	210/244 (86%) 216/244 (88%)
NOV30i	1..240 98..340	211/244 (86%) 217/244 (88%)

Further analysis of the NOV30a protein yielded the following properties shown in Table 30C.

Table 30C. Protein Sequence Properties NOV30a	
PSort analysis:	0.5500 probability located in endoplasmic reticulum (membrane); 0.1900 probability located in lysosome (lumen); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in outside
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV30a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several
5 homologous proteins shown in Table 30D.

Table 30D. Geneseq Results for NOV30a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV30a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB70549	Clone 16467945.0.85-S261.D protein sequence SEQ ID NO:82 - <i>Homo sapiens</i> , 546 aa. [WO200110902-A2, 15-FEB-2001]	32..450 65..483	417/419 (99%) 417/419 (99%)	0.0
AAB70547	Human PRO17 protein sequence SEQ ID NO:34 - <i>Homo sapiens</i> , 582 aa. [WO200110902-A2, 15-FEB-2001]	32..450 101..519	417/419 (99%) 417/419 (99%)	0.0
AAB80265	Human PRO334 protein - <i>Homo sapiens</i> , 509 aa. [WO200104311-A1, 18-JAN-2001]	36..450 88..473	383/415 (92%) 383/415 (92%)	0.0

AAU29049	Human PRO polypeptide sequence #26 - <i>Homo sapiens</i> , 509 aa. [WO200168848-A2, 20-SEP-2001]	36..450 88..473	383/415 (92%) 383/415 (92%)	0.0
AAY13397	Amino acid sequence of protein PRO334 - <i>Homo sapiens</i> , 509 aa. [WO9914328-A2, 25-MAR-1999]	36..450 88..473	383/415 (92%) 383/415 (92%)	0.0

In a BLAST search of public sequence databases, the NOV30a protein was found to have homology to the proteins shown in the BLASTP data in Table 30E.

Table 30E. Public BLASTP Results for NOV30a				
Protein Accession Number	Protein/Organism/Length	NOV30a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAC33425	Sequence 33 from Patent WO0110902 - <i>Homo sapiens</i> (Human), 582 aa.	32..450 101..519	417/419 (99%) 417/419 (99%)	0.0
Q91V88	POEM (NEPHRONECTIN short isoform) - <i>Mus musculus</i> (Mouse), 561 aa.	36..450 88..502	363/416 (87%) 386/416 (92%)	0.0
Q91ZD3	Nephronectin long isoform - <i>Mus musculus</i> (Mouse), 578 aa.	36..450 105..519	363/416 (87%) 386/416 (92%)	0.0
Q91XL5	Nephronectin - <i>Mus musculus</i> (Mouse), 592 aa.	38..450 121..533	362/414 (87%) 385/414 (92%)	0.0
Q923T5	Nephronectin - <i>Mus musculus</i> (Mouse), 609 aa.	38..450 138..550	362/414 (87%) 385/414 (92%)	0.0

PFam analysis predicts that the NOV30a protein contains the domains shown in

5 Table 30F.

Table 30F. Domain Analysis of NOV30a			
Pfam Domain	NOV30a Match Region	Identities/ Similarities for the Matched Region	Expect Value
EGF	41..75	15/47 (32%) 27/47 (57%)	0.84

EGF	81..115	10/47 (21%) 24/47 (51%)	0.034
EGF	166..201	12/47 (26%) 29/47 (62%)	4.9e-06

Fig. 1 shows that NOV30b (G51117-05) is expressed as about 66 kDa protein secreted by 293 cells.

Example 31.

5 The NOV31 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 31A.

Table 31A. NOV31 Sequence Analysis			
	SEQ ID NO: 123	3336 bp	
NOV31a, CG51264-01 DNA Sequence	CGCCGGTGGCTCGGCCGGCGGCCGGCGGCCGGCGGCCGGCGGCCGGCGTCTCTAC CTCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCCATCTGCTGTGGTTA TGGCCTGTCGCTGGAGCACAAAGAGTCTCCGCGGTGGAGGTCTGCGTTGCTCTTGCTTT TCCTCGCTGGGGTGTAAGCTTGTGGAGAGACTCCAGAGCAAATACGAGCACCAAGTGGCA TAATCACAAAGCCCAAGGCTGGCCTTCTGAATATCTTGCAAAAATCAACTGTAGCTGGTTCA TAAGGGCAAAACAGGCCAAATCATTACTATAAGTTTTCAGGATTTTGATATTCAAGGAT CCAGAAGGTGCAATTGGACTTGGTTGACAAATAGAAACATACAAGAATATTGAAAGTTTACA GAGCTTGTGGTTCCACAATTCCACCTCCGTATATCTCTTCACAAGACCACATCTGGATT GGTTTCATTTCGGATGACAACATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGA AATCTGAGGAACCAAATTGTGCTTGTGATCAGTTTCGTTGTGGTAATGGAAGTGTATAC CAGAAGCCTGGAATGCAATAACATGGATGAATGTGGAGATAGTTCGATGAAGAGATCT GTGCCAAAGAAGCAAACTCCCAACTGCTGCTGCTTTTCAACCTTGCTTACAACCAAGT TCCAGTGTTTATCCCGTTTTACCAAAGTTTACACTTGCCTCCCCGAATCTTTAAATGTG ATGGGAACATTGACTGCCTTGACCTAGGAGATGAGATAGACTGTGATGTGCCAACATGTG GGCAATGGCTAAATATTTTTATGGTACTTTTAATTCTCCCAATTATCCAGACTTTTATC CTCTCGGAAGCAATTGCACCTGGTTAATAGACACTGGTGATCACCGTAAAGTCATTTTAC GCTTCACTGACTTTAAACTTGTATGGTACTGGTTATGGTGATTATGTCAAAATATATGATG GATTAGAGGAGAATCCACACAAGCTTTTGGCGTGTGTTGACAGCTTTTGATTCTCATGCAC CTCTTACAGTTGTTTCTTCTCTGGACAGATAAGGGTACATTTTGTGCTGATAAAGTGA ATGCTGCAAGGGGATTTAATGCTACTTACCAAGTAGATGGGTCTGTTTGCCATGGGAAA TACCCTGTGGAGTAACTGGGGGTGTTATACTGAGCAGCAGCTGTGATGGGTATTTGGC ATTGCCCAAATGGAAGGGATGAAACCAATTGTACCATGTGCCAAGGAAGAATTTCCAT GTTCCCGAAATGGTGTCTGTTATCTCTCGTTCTGATCGCTGCAACTACCAGAATCATTGCC CAAATGGCTCAGATGAAAAAACTGCTTTTTTTGCCAACAGGAAATTTCCATTGTAAAA ACAATCGTTGTGTGTTTGAAGATTGGGTGTGTGATTCTCAAGATGACTGTGGTATGGCA GCGATGAAGAAAATTGCCAGTAATCGTGCCTACAAGAGTCATCACTGTCGCGTACATG GGAGCCTCATCTGTGGCCTGTACTCGTCATAGCATTTGGGATGACTTGTGAAGCTTTATT CTCTGAGAATGTTTGAAGAAGATCATTGAAACACAGTTGTCAAGAGTGGAAGCAGAAT TGTTAAGAAGAGAAGCTCCTCCTCGTATGGACAATTGATTGCTCAGGGTTTAATTCCAC CAGTTGAAGATTTTCTGTTTGTTCACCTAATCAGGCTTCTGTTTGGAAAAATCTGAGGC TAGCGGTACGATCTCAGCTTGGATTTACTTTCAGTCAGGCTTCCCTATGGCAGGCAGATCAA GCAACATTTGGAACCGTATTTTAAATTTGCAAGATCAGCTCACTTCTGAAGCTTATGGCTT TGGTCTCAGCAGATGGAGTAGAGTTGCCCTAGTCAGAGTACCAGTAGAGAACCTGAGA GAAATCATACTACAGAAGTTTGTGTTTCCGTGGAGTCTGATGATACAGACACAGAAAATG AGAGAAGAGATATGGCAGGAGCATCTGGTGGGGTTGCAGCTCCTTTGCCTCAAAAAGTCC CTCCCAACCGGCAGTAGAAGCGACAGTAGGAGCATGTGCAAGTTTCTCAACTCAGAGTA CCCGAGGTGGTCATGCAGATAATTGGAAGGGATGTGACAAGTGTGGAACCCCAAGTGTGA GTCCAGCAGCTCACCAGGTTACAAGTGCACTCAGTCGTATGACTCAGGGGCTACGCTGGG		

[illegible]

	CCCGAATGGTGCTCTTTACTCTCGTTCTGATCGCTGCAACTACCAAGTACCCATTGACCAAA ATGGCAAACAGAACCCATCTACTTGGTAAGTAGCATTAAATCCCCCTTGCAGCATTAC		
	ORF Start: ATG at 120		ORF Stop: TAA at 1467
	SEQ ID NO: 126	449 aa	MW at 50654.0kD
NOV31b, CG51264-03 Protein Sequence	MACRWSTKESPRWRSALLLLFLAGVYNGALAEHSENVHISGVSTACGETPEQIRAPSGI ITSPGWPSBYPAKINCSWFIRANPGEIITISFQDFDIQGSRRCNLDWLTIIETKNIESYR ACGSTIPPPYISSQDHIWIRFHSDDNISRKGFRLAYFSGKSEEPNACDQFRGNGKCIPE EAWKCNMDECGDSSDEEICAKEANPPTAAAFQPCAYNQFQLSRFTKVYTCLEPESLKCD GNIDCLDLGDEIDCDVPTCCQWLKYFYGTFFNSPNYPDFYPPGSNCTWLIDTGDHRKVL FTDFKLDGTGYGDYVKIYDGLEENPHKLLRLVLTAFDASHPLTVVSSSGQIRVHFCADKVN AARGFNATYQVDGFCLPWEIPCGGNWGCYTEQQRDGYWHCPNGRDETNCTMCQKEEFP SRNGVCYPRSDRCNYQNHCPNGKQNPSTW		
	SEQ ID NO: 127	1441 bp	
NOV31c, CG51264-04 DNA Sequence	CGCCGGTGGCTCGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCTAC CTCCAGCTTCTCCTCCCTCCTCCTCGTCTCCTCCTCTCTCTCTCCATCTGCTGTGGTTA TGGCCTGTGCTGGAGCACAAAAGAGTCTCCGCGGTGGAGGTCTGCCTTGCTCTTGCTTT TCCTCGCTGGGGTGTACGCTTGTGGAGAGACTCCAGAGCAAAATACGAGACCAAGTGGCA TAATCACAAGCCAGGCTGGCCTTCTGAATATCTGCAAAAATCAACTGTAGCTGGTTCA TAAGGGCAAACCCAGGCGAAATCATTACTATAAGTTTTAGGATTTTGATATTCAAGGAT CCAGAAGGTGCAATTTGGACTGGTTGACAATAGAAACATACAAGAATATTGAAAGTTACA GAGCTTGTGGTTCACAATCCACCTCCGTATATCTCTTACAAGACCACATCTGGATTA GGTTCATTCCGATGACACATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGA AATCTGAGGAACCAAATTTGCTTGTGATCAGTTTCGTTGGTAAAGTGAAGGTATAC CAGAAGCCTGGAATGTAATAACATGGATGAATGTGGAGATAGTTCGGATGAAGAGATCT GTGCCAAAGAAGCAAATCTCCAAGTCTGCTGCTTTTCAACCTGTGCTTACAACCAAGT TCCAGTGTATATCCCGTTTACCAAAGTTTACACTTGCCTCCCCGAATCTTTAAATGTG ATGGGAACATTGACTGCCTTTGACCTAGGAGATGAGATAGACTGTGATGTCGAACATGTG GGCAATGGCTAAAATTTTATGTTACTTTTAACTTCCAATCTGCAAGTCTGGACTTTTAC CTCCTGGAAGCAATTGCACCTGGTTAATAGACACTGGTGATCACCCTAAAGTCATTTTAC GCTTCACTGACTTTAACTTGATGGTACTGGTTATGGTGATTATGTCAAATATATGATG GATTAGAGGAGAAATCCACAAAGCTTTTGCCTGTGTTGACAGCTTTTGATTCTCATGCAC CTCTTACAGTTGTTTCTTCTTGGACAGATAAGGGTACATTTTGTGCTGATAAAGTGA ATGCTGCAAGGGGATTTAATGCTACTTACCAAGTAGATGGGTTCTGTTGCTGCAAGGAAA TACCTGTGGAGGTAATCGGGGTGTTATACTGAGCAGCAGCGTCGTGATGGGTATTGGC ATTGCCCAAATGGAAGGGATGAACCAATTGTACCATGTGCCAGAAGGAAGAAATTTCCAT GTTCCCGAAATGGTGTCTGCTATCTCTGCTCTGATCGCTGCAACTACCAGAATCATTGCC CAAATGGCAAACAGAACCCATCTACTTGGTAAGTAGCATTAAATCCCCCTTGCAGCATTCA C		
	ORF Start: ATG at 120		ORF Stop: TAA at 1410
	SEQ ID NO: 128	430 aa	MW at 48793.0kD
NOV31c, CG51264-04 Protein Sequence	MACRWSTKESPRWRSALLLLFLAGVYACGETPEQIRAPSGIITSPGWPSBYPAKINCSWF IRANPGEIITISFQDFDIQGSRRCNLDWLTIIETKNIESYRACGSTIPPPYISSQDHIWI RFHSDDNISRKGFRLAYFSGKSEEPNACDQFRGNGKCIPEAWKCNMDECGDSSDEEICA KEANPPTAAAFQPCAYNQFQLSRFTKVYTCLEPESLKCDGNIDCLDLGDEIDCDVPTC GQWLKYFYGTFFNSPNYPDFYPPGSNCTWLIDTGDHRKVLRFDTDFKLDGTGYGDYVKIYD GLEENPHKLLRLVLTAFDASHPLTVVSSSGQIRVHFCADKVNAAARGFNATYQVDGFCLPWE IPCGGNWCYTEQQRDGYWHCPNGRDETNCTMCQKEEFPSCRNGVCYPRSDRCNYQNHCP NGKQNPSTW		
	SEQ ID NO: 129	3021 bp	
NOV31d, CG51264-06 DNA Sequence	CTCCTCCTCCGCTCCTCCTCCTCCTCCTCATCTGCTGTGGTTATGGCCTGTGCTGGAGC ACAAAGAGTCTCCGCGGTGGAGGTCTGCGTTGCTCTTGCTTTCTCCTCGCTGGGGTGTAC GCTTGTGGAGAGACTCCAGAGCAAAATACGAGCACCAAGTGGCATAATACAAGCCAGGC TGGCCTTCTGAATATCTGCAAAAATCAACTGTAGCTGGTTTATAAGGGCAAACCCAGGC GAAATCATTACTATAAGTTTTCAGGATTTTGATATTCAAGGATCCAGAAGGTGCAATTTG GACTGTTGCAATAGAAACATACAAGAATATTGAAAGTTACAGAGCTGTGGTTCACACA ATTCCACCTCCGTATATCTCTTACAAGACCACATCTGGATTAGGTTTTCATTCCGGATGAC		

	AACATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGAAATCTGAGGAACCAAAT TGTGCTTGTGATCAGTTTCGTTGTGGTAATGGAAAGTGTATACCAGAAGCCTGGAAATGT AATAACATGGATGAATGTGGAGATAGTTCCGATGAAGAGATCTGTGCCAAAGAAGCAAAT CCTCCAACTGCTGCTGCTTTTCAACCCCTGTGCTTACAACCAAGTTCAGTGTTTATCCCGT TTTACCAAAGTTTACACTTGCCCTCCCGAATCTTTAAATGTGATGGGAACATTGACTGC CTTGACCTAGGAGATGAGATAGACTGTGATGTGCCAACATGTGGCAATGGCTAAATAT TTTTATGGTACTTTTAATTCTCCAATTATCCAGACTTTTATCCTCCTGGAAGCAATTGC ACCTGGTTAATAGACACTGGTGATCACCGTAAAGTCATTTTACGCTTCACTGACTTTAAA CTTGATGGTACTGGTTATGGTGATTATGTCAAATATATGATGGATTAGAGGAGAATCCA CACAAGCTTTTGCGTGTGTTGACAGCTTTTGATTCTCATGCACCTCTTACAGTTGTTTCT TCTTCTGGACAGATAAGGGTACATTTTGTGCTGATAAAGTGAATGCTGCAAGGGGATTT AATGCTACTTACCAAGTAGATGGGTTCTGTTTGCCATGGGAAATACCCTGTGGAGGTAAC TGGGGGTGTTATACTGAGCAGCAGCGTTGTGATGGGTATTGGCATTGCCCAAATGGAAGG GATGAAACCAATTGTACCATGTGCCAGAAGGAAGAATTTCCATGTTCCCGAAATGGTGTC TGTTATCCTCGTTCTGATCGCTGCAACTACCAGAATCATTGCCCAAATGGCTCAGATGAA AAAAACTGCTTTTGTGCCAACAGGAAATTTCCATTGTAAAAACAATCGTTGTGTTT GAAAGTTGGGTGTGTGATTCTCAAGATGACTGTGGTGATGGCAGCGATGAAGAAAATTGC CCAGTAATCGTGCCTACAAGAGTCATCACTGCTGCCGTCATAGGGAGCCTCATCTGTGGC CTGTTACTCGTCATAGCATTGGGATGTACTTGTAAAGCTTTATCTCTGAGAATGTTTGAA AGAAGATCATTTGAAACACAGTTGTCAAGAGTGAAGCAGAATTGTTAAGAAGAGAAGCT CCTCCCTCGTATGACAATTGATTGCTCAGGGTTAATTCCACCAGTTGAAGATTTTCCCT GTTTGTTCACCTAATCAGGCTTCTGTTTGGAAAATCTGAGGCTAGCGTACGATCTCAG CTTGATTTACTTCACTCAGGCTTCTATGGCAGGCAGATCAAGCAACATTGGAACCGT ATTTTAAATTTTGCAAGATCACGTCATTCTGGGTATTGGCTTTGGTCTCAGCAGATGGA GATGAGGTTGTCCCTAGTCAGAGTACCAGTAGAGAACCTGAGAGAAATCATACTCACAGA AGTTTGTTTTCCGTGGAGTCTGATGATACAGACACAGAAAATGAGAGAAGAGATATGGCA GGAGCATCTGGTGGGTTGACGCTCCTTGCCTCAAAAAGTCCCTCCCACAACGGCAGTG GAAGCGACAGTAGGAGCATGTGCAAGTTCCTCAACTCAGAGTACCCGAGGTGGTTCATGCA GATAATGGAAGGGATGTGACAAGTGTGAACCCCAAGTGTGAGTCCAGCACGTCACCAG CTTACAAGTGCCTCAGTCGTATGACTCAGGGGCTACGCTGGGTACGTTTTACATTAGGA CGATCAAGTTCCTAAGTCAGAACCAGAGTCCTTTGAGACAACTTGATAATGGGGTAAGT GGAAGAGAAGATGATGATGATGTTGAAATGCTAATTCCAATTTCTGATGGATCTTCAGAC TTTGATGTGAATGACTGCTCCAGACCTCTTCTTGATCTTGCCTCAGATCAAGGACAAGGG CTTAGACAACCATATAATGCAACAAATCCTGGAGTAAGGCCAAGTAATCGAGATGGCCCC TGTGAGCGCTGTGTTATTGTCCACACTGCCAGATACCAGACACTTGCTTAGAAGTAACA CTGAAAAACGAAACGAGTGATGATGAGGCTTTGTTACTTTGTAGGTACGAATCACATAA GGGAGATTGTATACAAGTTGGAGCAATATCCATTATTATTGTAACCTTACAGTTAAA CTAGTTTTAGTTTAAAAAGAAAAATGCAGGGTGATTCTTATTATTATATGTTAGCCTG CATGGTTAAATTCGACAACTTGTAACCTATGAACCTAGAGTTTACTATTTTAGCAGCTA AAAATGCATCACATATTGCATATTGTTCAATAATGGTCCTTTCATTTGTTTCTGATTGTT TTCATCCTGATACTGTAGTTCACTGTAGAAATGTGGCTGCTGAACTCATTGATTGTCA TTTTATCTATCTATGTTAAATGGTTGTTTTTACAAAATAATACCTATTTTTAATTGA AACGTTTATGCTTTTGCCAAGCACATCTTGTAACCTAATATAGCTAGATGTTAAGTTGT TAATGTACCAAAAAAAAAA		
	ORF Start: ATG at 43		ORF Stop: TAG at 2563
	SEQ ID NO: 130	840 aa	MW at 93121.8kD
NOV31d, CG51264-06 Protein Sequence	MACRWSTKESPRWRSALLLFLAGVYACGETPEQIRAPSGIITSPGWSEYPKINCSWF IRANPGEIITISFQDFDIQGSRRCLNDWLTIIETKNIESYRACGSTIPPPYISSQDHIWI RFHSDDNISRKFRLAYFSGKSEEPNCACDQFRCGNGKCIPEAWKCNMDECGDSSDEEI CAKEANPPTAAAFQPCAYNQFQCLSRFTKVYTCLPESLKCDGNIDCLDLGDEIDCDVPTC GQWLKYFYGTFNSPNYPDFYPPGSNCTWLIDTGDHRKVLIRFTDFKLDGTGYGDYVKIYD GLEENPHKLLRLVAFDASHAPLTVVSSSGQIRVHFCADKVNAARGFNATYQVDGFLPWE IPCGGNWGCYTEQQRCDGYWHCPNGRDETNTMCQKEEFPCSRNGVCYPRSDRCNYQNHC PNGSDEKNCFQCQPGNFHCKNNRCVFESWVCDSDQDCGDSDEENCPIVPTRVITA AVI GSLICGLLLVIALGCTCKLYSLRMFERRSFETQLSRVEAELLRREAPPSYGLIAQGLIP PVEDFPVCSPNQASVLENLRLAVRSQLGFTSVRLPMAGRSSNIWNRI FNFARSRHSGSLA LVSADGDEVVPSQSTSREPRNHTHRSLSVESDDTDENERRDMAGASGGVAAPLPQKV PPTTAVEATVGACASSSTQSTRGGHADNGRDVTSVEPPSVSPARHQLTSALSMTQGLRW		

	VRFTLGRSSSLSONQSPRLQLDNGVSGREDDDDVEMLIPIISDGSSDFDVNDCSRPLDLA SDQGQGLRQPYNATNPGRVPSNRDGPCECGRGIVHTAQIPDTCLEVTLKNETSDDEALLLC		
	SEQ ID NO: 131	3012 bp	
NOV31e, CG51264-07 DNA Sequence	CTCCTCCTCCGTCTCCTCCTCTCTCTCATCTGCTGTGGTTATGGCCTGTGCTGGAGC ACAAAAGAGTCTCCGCGGTGGAGGTCTGCGTTGCTCTTGCTTTCTCGCTGGGGTGAC GCTGTGAGAACTCAACAATACAGCACAAAGTGGCATAATCACAAGCCCAGGCTGGCCTTCT GAATATCCTGCAAAAATCAACTGTAGCTGGTTTCATAAGGGCAAACCCAGGCGAAATCATT ACTATAAGTTTTTCAAGATTTTGATATTCAAGGATCCAGAAGGTGCAATTTGGACTGGTTG ACAATAGAAACATACAAGAATATTGAAAGTTACAGAGCTTGTGGTTCCACAATTCACCT CCGTATATCTCTTCAAGACCACATCTGGATTAGGTTTCATTCCGATGACAACATCTCT AGAAAGGGTTTCAGACTGGCATATCTTTCAAGCAAATCTGAGGAACCAAATTTGTGCTTGT GATCAGTTTCGTTGTGGTAATGGAAGTGTATACCAGAAGCTGGAATGTAATAACATG GATGAATGTGGAGATAGTTCGGATGAAGAGATCTGTGCCAAGAAGCAAATCTCCAAC GCTGCTGCTTTTCAACCTGTGCTTACAACCAAGTCCAGTGTATTCCCGTTTACCAAA GTTTACACTTGCTCCCGAATCTTTAAATGTGATGGGAACATTGACTGCCTTGACCTA GGAGATGAGATAGACTGTGATGTGCCAACATGTGGGCAATGGCTAAATATTTTATGGT ACTTTTAATTCTCCAATTATCCAGACTTTTATCCTCCTGGAAGCAATTGCACCTGGTTA ATAGACACTGGTGATCACCGTAAAGTCATTTTACGCTTCACTGACTTTAAACTTGATGGT ACTGGTTATGGTGATTATGTCAAATATATGATGGATTAGAGGAGAATCCACACAAGCTT TTGCGTGTGTTGACAGCTTTGATTCTCATGCACCTCTTACAGTTGTTTCTTCTCTGGA CAGATAAGGGTACATTTTGTGCTGATAAAGTGAATGCTGCAAGGGGATTAATGCTACT TACCAAGTAGATGGGTTCTGTTTGCCATGGGAAATACCCTGTGGAGGTAACGGGGGTGT TATACTGAGCAGCAGCGTTGTGATGGGTATTGGCATTGCCCAAATGGAAGGGATGAAACC AATTGTACCATGTGCCAGAAGGAAGAATTTCCATGTTCCGAAATGGTGTCTGTATTCTCT CGTTCTGATCGCTGCAACTACCAGAATCATTGCCCAAATGGCTCAGATGAAAAAAGCTGC TTTTTTGCCAACCAGGAAATTTCCATTGTAAAAACAATCGTTGTGTGTTTGAAGTTGG GTGTGTGATTCTCAAGATGACTGTGGTGATGGCAGCGATGAAGAAAATTGCCAGTAATC GTGCCTACAAGAGTCATCACTGCTGCCGTATAGGGAGCCTCATCTGTGGCCTGTACTC GTCATAGCATTGGGATGTACTTGTAAAGCTTTATTCTCTGAGAATGTTTGAAGAAGATCA TTTGAACACAGTTGTCAAGAGTGGGAAGCAGAATGTTAAGAAGAGAAGCTCCTCCCTCG TATGGACAATTGATTGCTCAGGGTTTAATTCACCAGTTGAAGATTTTCTGTTTGTTC CCTAATCAGGCTTCTGTTTGGAAAATCTGAGGCTAGCGGTACGATCTCAGCTTGATTT ACTTCAGTCAGGCTTCTATGGCAGGCAGATCAAGCAACATTTGGAACCGTATTTTAAT TTTGCAAGATCACGTCATTCTGGGTCTTGGCTTTGGTCTCAGCAGATGGAGATGAGGT GTCCCTAGTCAGAGTACCAGTAGAGAACCTGAGAGAAATCATACTCACAGAAGTTGTGTT TCCGTGGAGTCTGATGATACAGACACAGAAAATGAGAGAAGAGATATGGCAGGAGCATCT GGTGGGGTTGCAGCTCCTTTGCCCTCAAAAAGTCCCTCCACAACGGCAGTGGAAGCGACA GTAGGAGCATGTGCAAGTTCTCAACTCAGAGTACCCGAGGTGGTCATGCAGATAATGGA AGGGATGTGACAAGTGTGGAACCCCAAGTGTGAGTCCAGCACGTCACCAGCTTACAAGT GCACTCAGTCGATGACTCAGGGGTACGCTGGGTACGTTTACATTAGGACGATCAAGT TCCCTAAGTCAGAACCAGAGTCCTTTGAGACAACCTGATAATGGGGTAAAGTGAAGAGAA GATGATGATGATGTTGAAATGCTAATCCAATTTCTGATGGATCTTCAGACTTTGATGTG AATGACTGCTCCAGACCTCTTCTGATCTTGCCCTCAGATCAAGGACAAGGGCTTAGACAA CCATATAATGCAACAAATCCTGGAGTAAGGCCAAGTAATCGAGATGGCCCTGTGAGCGC TGTGGTATTGTCCACACTGCCAGATACCAGACACTTGCTTAGAAGTAACACTGAAAAAC GAAACGAGTGATGATGAGGCTTTGTACTTTGTAGGTACGAATCACATAAGGGAGATTG TATACAAGTTGGAGCAATATCCATTTATTATTTGTAACCTTACAGTTAACTAGTTTTA GTTTAAAAAGAAAAATGCAGGGTGATTTCTTATTATTATATGTTAGCTGCATGGTTAA ATTCGACAACCTGTAACCTATGAACCTAGAGTTTACTATTTAGCAGCTAAAAATGCAT CACATATTGCATATTGTTCAATAATGGTCCTTTTCAATTTGTTTCTGATTGTTTTCATCCTG ATACTGTAGTTCACTGTAGAAATGTGGCTGCTGAAACTCATTGTGATTTGCTATTATCT ATCCTATGTTAAATGGTTGTTTTTACAAAATAATACCTTATTTAATTGAAACGTTTAT GCTTTTGCCAAGCACATCTGTAACTTAATATAGCTAGATGTTAAGGTGTTAATGTACC AAAAAAAAAAAA		
	ORF Start: ATG at 43		ORF Stop: TAG at 2554
	SEQ ID NO: 132	837 aa	MW at 92869.5kD
NOV31e, CG51264-07	MACRWSTKESPRWRSALLLFLAGVYAVRTQQYSTSGIITSPGWPSEYPAKINCSWFIRA NPGEIITISFQDFDIQGSRRCNLDWLTIIETYKNIESYRACGSTIPPPYISSQDHIWIRFH		

Protein Sequence	<p>SDNNISIRKGRFLAYLSGKSEEPNCADQFRCGNGKCIPEAWKCNMDECGDSSDEEIKAK EANPPTAAAFQPCAYNQFCLSRFTKVYTCPLPESLKCNDICDLDLGDEIDCDVPTCGQW LKYFYGTFNSPNYPDFYPPGSNCTWLIDTGDHRKVI LRFTDFKLDGTGYGDYVKIYDGL ENPHKLLRVLTAFD SHAPLTVVSSSGQIRVHFCADKVNAAARGFNATYQVDGFCPLPWEIPC GGNWGCYTEQQRCDGYWHCPNGRDETNTMCQKEEFPCSRNGVCYPRSDRCNYQNHCPNG SDEKNCFFCQPGNFHCKNNRCVFESWVCDSDQDCGDSDEENCPVIVPTRVITAAVIGSL ICGLLLVIALGCTCKLYSLRMFFERSFETQLSRVEALLRREAPPYGLIAQGLIPPE DFPVCSNPQASVLENLRLAVRSQLGFTSVRLPMAGRSSNIWNRI FNFARSRHSGSLALVS ADGDEVVPSQSTSREPERNHTHRSLSFVSESDDTDTENERRDMAGASGGVAAPLPQKVPP TAVEATVGACASSSTQSTRGGHADNGRDVTSEPPSPS PARHQLTSALSMTQGLRWVRI TLRSSSSLSQNSQPLRQLDNGVSGREDDDDVEMLIPISDGSSDFDVNDCSRPLDLASDQ GQGLRQPNYATNPGVRPSNRDGPCECRGIVHTAQIPDTCLEVTLKNETSDDEALLLC</p>		
	SEQ ID NO: 133	1441 bp	
NOV31f, CG51264-02 DNA Sequence	<p>CGCCGGTGGCTCGGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGTCTAC CTCAGCTTCTCCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCATCTGCTGTGTTA TGGCCTGTCGCTGGAGCACAAAGAGTCTCCGCGGTGGAGGTCTGCGTTGCTCTTGCTTT TCCTCGCTGGGGTGACGCTTGTGGAGAGACTCCAGAGCAAATACGAGACCAAGTGGCA TAATCACAAGCCCAGGCTGGCCTTCTGAATATCCTGCAAAATCAACTGTAGCTGGTTCA TAAGGGCAAACCCAGGCGAAATCATTACTATAAGTTTTCAGGATTTGATATTCAAGGAT CCAGAAGGTGCAATTTGGACTGGTTGACAATAGAAACATACAAGAATATTGAAAGTTACA GAGCTTGTGGTTCACAATTCACCTCCGTATATCTCTTCACAAGACCACATCTGGATTA GGTTTCATTTCGATGACAACATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTAGGGA AATCTGAGGAACCAAATTTGTGCTTGTGATCAGTTTCGTGTGGTAATGGAAAGTGATAC CAGAAGCTTGGAAATGTAATAACATGGATGAATGTGAGATAGTTCCGATGAAGAGATCT GTGCCAAAGAAGCAAATCCTCCAACCTGCTGCTGCTTTTCAACCCTGTGCTTACAACCAGT TCCAGTGTTCATCCGTTTACCAAAGTTTACACTTGCTCCCGAATCTTTAAATGTG ATGGGAACATTGACTGCCTTGACCTAGGAGATGAGATAGACTGTGATGTGCCAACATGTG GGCAATGGCTAAAATATTTTATGGTACTTTTAATTCCTCAATATCAGAGCTTTTATC CTCTGGAAGCAATGGCACTGGTTAATAGACACTGGTGATCCGTAAGACTTTTATC GCTTCACTGACTTTAACTTGATGGTACTGGTTATGGTGATTATGTCAAATATATGATG GATTAGAGGAGAATCCACACAAGCTTTTGGCTGTGTGACAGCTTTTGATTCTCATGCAC CTCTTACAGTTGTTTCTTCTTCTGACAGATAAGGGTACATTTTGTGCTGATAAAGTGA ATGCTGCAAGGGGATTAATGCTACTTACCAAGTAGATGGGTCTGTTGCCATGGGAAA TACCTTGAGGATTAAGTGGGGTGTTATACTGAGCAGCAGCGCTGATGGGTATTTGGC ATTGCCCAAATGGAAGGGATGAAACCAATTGTACCATGTGCCAGAAGGAAGAAATTCAT GTTCCCGAAATGGTGTCTGCTATCCTCGCTCTGATCGCTGCAACTACCAGAATCATTGCC CAAATGGCAAACAGAACCCATCTACTTGGTAAGTAGCATTAATCCCCCTTCAGCATTCAC C</p>		
	ORF Start: at 3		ORF Stop: TAA at 1410
	SEQ ID NO: 134	469 aa	MW at 53338.2kD
NOV31f, CG51264-02 Protein Sequence	<p>PVARRRRRRRRRRRRRLPPASPPSSSVSSSLSPSAVVMACRWSTKESPRWRSALLLF LAGVYACGETPEQIRAPSGIITSPGPWSEYPKINCWSFIRANPGEIITISFQDFDIQGS RRCNLDWTIETKYNIERSYACGSTIPPPYISSQDIWIRFSDNNISIRKGRFLAYFSGK SEEPNCADQFRCGNGKCIPEAWKCNMDECGDSSDEEIKAKEANPPTAAAFQPCAYNQF QCLSRFTKVYTCPLPESLKCNDICDLDLGDEIDCDVPTCGQWLKYFYGTFNSPNYPDFY PGSNCTWLIDTGDHRKVI LRFTDFKLDGTGYGDYVKIYDGLEENPHKLLRVLTAFD SHAP LTVVSSSGQIRVHFCADKVNAAARGFNATYQVDGFCPLPWEIPCGGNWGCYTEQQRDGYWH CPNGRDETNTMCQKEEFPCSRNGVCYPRSDRCNYQNHCPNGKQNPSTW</p>		
	SEQ ID NO: 135	3078 bp	
NOV31g, CG51264-05 DNA Sequence	<p>CTCCTCCTCCGCTCCTCCTCCTCCTCCTCATCTGCTGTGGTTATGGCCTGTGCTGGAGC ACAAAGAGTCTCCGCGGTGGAGTCTGCGTTGCTCTTGCTTTTCTCGCTGGGGTGAC GGAAATGGTGCTCTTGAGAACATTCTGAAATGTGCATATTTAGGAGTGTCAACTGCT TGTGGAGAGACTCCAGAGCAAATACGAGCACCAAGTGGCATAATACAAGCCAGGCTGG CTTCTGAATATCCTGCAAAATCAACTGTAGCTGGTTCATAAGGGCAAACCCAGGCGAA ATCATTACTATAAGTTTTCAGGATTTTGATATTCAAGGATCCAGAAGGTGCAATTTGGAC TGTTTGACAATAAGAACATAAGAAATATTGAAAGTACAGAGTCTGTGGTTCACAATT CCACCTCCGTATATCTTTCACAAGACCACATCTGGATTAGGTTTCATTTCGATGACAA</p>		

	<p>ATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGAAATCTGAGGAACCAAATGT GCTTGTGATCAGTTTCGTTGTGGTAATGGAAAGTGTATACCAGAAGCCTGGAAATGCAAT AACATGGATGAATGTGGAGATAGTTCCGATGAAGAGATCTGTGCCAAAGAGCAAATCCT CCAAGTGTCTGCTGCTTTTCAACCTGTGCTTACAACCAAGTTCAGTGTATATCCCGTTT ACCAAAGTTTACACTTGCCTCCCCGAATCTTTAAATGTGATGGGAACATTGACTGCCTT GACCTAGGAGATGAGATAGACTGTGATGTGCCAACATGTGGGCAATGGCTAAAATATTTT TATGGTACTTTTAATTCTCCCAATTATCCAGACTTTTATCCTCCTGGAAGCAATTGCACC TGGTTAATAGACACTGGTGATCACCGTAAAGTCATTTACGCTTCACTGACTTTAACTT GATGGTACTGGTTATGGTGATTATGTCAAAATATATGATGGATTAGAGGAGAATCCACAC AAGCTTTTGGTGTGTTGACAGCTTTTGATTCTCATGCACCTCTTACAGTTGTTTCTTCT TCTGGACAGATAAGGGTACATTTTGTGTGATGATAAAGTGAATGCTGCAAGGGGATTTAAT GCTACTTACCAAGTAGATGGGTTCTGTTTGCCATGGGAAATACCTGTGGAGGTAAGTGG GGGTGTTTACTGAGCAGCAGCGTTGTGATGGGTATTGGCATTGCCAAATGGAAGGGAT GAAACCAATGTACCATGTGCCAGAAGGAAGAATTTCCATGTTCCCGAAATGGTGTCTGT TATCCTCGTTCTGATCGCTGCAACTACCAGAATCATTGCCCAATGGCTCAGATGAAAAA AACTGCTTTTGTGCCAACAGGAAATTTCCATTGTAAAAACAATCGTTGTGTGTTGAA AGTTGGGTGTGTGATTCTCAAGATGACTGTGGTGATGGCAGCGATGAAGAAAATTGCCCA GTAATCGTGCCTACAAGAGTCATCACTGCTGCCGTCATAGGGAGCCTCATCTGTGGCCTG TTACTCGTCATAGCATTGGGATGTACTTGAAGCTTTATTCTCTGAGAATGTTTGAAGA AGATCATTGAAACACAGTTGTCAAGAGTGAAGCAGAATTGTTAAGAAGAGAAGCTCCT CCCTCGTATGGACAATTGATTGCTCAGGGTTAATTCCACCAGTTGAAGATTTTCCCTGTT TGTTACCTAATCAGGCTTCTGTTTGGAAAATCTGAGGCTAGCGGTACGATCTCAGCTT GGATTTACTTCAGTCAGGCTTCTATGGCAGGCAGATCAAGCAACATTTGGAACCGTATT TTTAATTTTGCAAGATCACGTCATTCTGGGTCAATTGGCTTTGGTCTCAGCAGATGGAGAT GAGGTTGTCCCTAGTCAGAGTACCAGTAGAGAACCTGAGAGAAATCATACTCACAGAAGT TTGTTTTCCGTGGAGTCTGATGATACAGACACAGAAAATGAGAGAAGAGATATGGCAGGA GCATCTGGTGGGGTTGCAGCTCCTTTGCCCTCAAAAAGTCCCTCCCAACGGCAGTGGAA GCGACAGTAGGAGCATGTGCAAGTTCCTCAACTCAGAGTACCCGAGGTGGTCATGCAGAT AATGGAAGGGATGTGACAAGTGTGGAACCCCCAAGTGTGAGTCCAGCACGTCACCAGCTT ACAAGTGCACCTCAGTCGTATGACTCAGGGGCTACGCTGGGTACGTTTTACATTAGGACGA TCAAGTTCCTAAGTCAGAACCAGAGTCCTTTGAGACAACTTGATAATGGGGTAAGTGA AGAGAAGATGATGATGATGTTGAAATGCTAATTCGAATTTCTGATGGATCTTCAGACTTT GATGTGAATGACTGCTCCAGACCTCTTCTTGATCTGCCTCAGATCAAGGACAAGGGCTT AGACAACCATATAATGCAACAAATCCTGGAGTAAGGCCAAGTAATCGAGATGGCCCCCTGT GAGCGCTGTGGTATTGTCCCACTGCCAGATACCAGACACTTGCTTAGAAGTAACACTG AAAAACGAAACGAGTGATGATGAGGCTTTGTTACTTTGTTAGGTACGAATCACATAAGGG AGATTGTATACAAGTTGGAGCAATATCCATTTATTATTTGTAACCTTTACAGTTAACTA GTTTTAGTTTAAAAAGAAAAATGCAGGGTGATTTCTTATTATTATATGTTAGCCTGCAT GGTTAAATTCGACAACTTGTAACTCTATGAACCTAGAGTTTACTATTTTAGCAGCTAAAA ATGCATCACATATTGCATATTGTTCAATAATGGTCCTTTTCATTGTTTCTGATTGTTTC ATCCTGATACTGTAGTTCACTGTAGAAATGTGGCTGCTGAAACTCATTGTATTGTCAATT TTATCTATCCTATGTTAAATGGTTTGTTTTACAAAATAATACCTTATTTTAAATTGAAC GTTTATGCTTTTGGCAAGCACATCTTGTAACTTAATATAGCTAGATGTTAAGGTTGTTAA TGTACCAAAAAAAAAAAAA</p>
	<p>ORF Start: ATG at 43</p>
	<p>ORF Stop: TAG at 2620</p>
	<p>SEQ ID NO: 136</p>
	<p>859 aa</p>
	<p>MW at 94982.7kD</p>
<p>NOV3lg, CG51264-05 Protein Sequence</p>	<p>MACRWSTKESPRWRSALLLFLAGVYNGALAEHSENVHISGVSTACGETPEQIRAPSGI ITSPGWSEYPKINCSWFIRANPGEIITISFQDFDIQGSRRCNLDWLTETYKNIESYR ACGSTIPPPYISSQDHIWIRFHSDDNISRKGFRLAYFSGKSEEPNCACDFRCGNGKCI EAWKCNMDECGDSSDEEICAKEANPTAAAFQPCAYNQFQCLSRFTKVYTCLPESLKCD GNIDCLDLGDEIDCDVPTCGQWLKIFYGTFNPNYPDFYPPGSNCTWLIDTGDHRKVLIR FTDFKLDGTGYDYVKIYDGLBENPHKLLRVLTAFDHAPLTVVSSGQIRVHFCADKVN AARGFNATYQVDGFCLPWEIPCGGNWGCYTEQQRCDGYWHCPNGRDETNTMCQKEEPPC SRNGVCYPRSDRCNYQNHCPNGSDEKNCFFCQPGNFHCKNNRCVFESWVCSQDDCGDGS DEENCPIVPTRVITAIVIGSLICGLLLVIALGCTCKLYSLRMFERRSFETQLSRVBAEL LRREAPPSYQGLIAQGLIPVEDFPVCSPNQASVLENLRLAVRSQLGFTSVRLPMAGRSS NIWNRI FNFARSRHSGSLALVSADGEVVPVSQSTSREPERNHTRSLFSVESDDTDENE RRDMAGASGGVAAPLPQKVPPTTAVEATVGACASSSTQSTRGGHADNGRDVTSVEPPSVS</p>

	PARHQLTSALSRTQGLRWVRFTLGRSSSLSQNQSPLRQLDNGVSGREDDDDVEMLIPI DGSSDFDVNDCSRPLDLASDQGGQLRQYPYNATNPGVRPSNRDGPCERCGIVHTAQIPDT CLEVTLKNETSDDEALLLC	
	SEQ ID NO: 137	1389 bp
NOV31h, CG51264-08 DNA Sequence	AATGGTGCTCTTGCAGAACATTCTGAAAATGTGCATATTTCAAGGAGTGTCAACTGCTTGT GGAGAGACTCCAGAGCAAATACGAGACCAAGTGGCATAATCACAAGCCCAGGCTGGCCT TCTGAATATCCTGCAAAAACCAACTGTAGCTGGTTCATAAGGGCAAACCCAGGCGAAATC ATTACTATAAGTTTTCAGGATTTTGATATTCAAGGATCCAGAAGGTGCAATTTGGACTGG TTGACAATAGAAACATACAAGAATATTGAAAGTTACAGAGCTTGTGGTTCACACAATTTCCA CCTCCGTATATCTCTTCACAAGACCACATCTGGATTAGGTTTCATTCCGATGACAACATC TCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGAAATCTGAGGAACCAAATTTGTGCT TGTGATCAGTTTCGTTGTGGTAATGGAAAGTGTATACCAGAAGCCTGGAAATGTAATAAC ATGGATGAATGTGGAGATAGTTCCGATGAAGAGATCTGTGCCAAAGAAGCAAATCCTCCA ACTGCTGCTGCTTTTTCAACCCTGTGCTTACAACCAAGTCCAGTGTTTATCCCGTTTTTACC AAAGTTTACACTTGCTCCCCGAATCTTTAAAATGTGATGGGAACATTGACTGCCTTGAC CTAGGAGATGAGATAGACTGTGATGTGCCAACATGTGGGCAATGGCTAAAATATTTTTAT GGTACTTTTTAATTCTCCCAATTATCCAGACTTTTATCCTCCTGGAAGCAATTGCACCTGG TTAATAGACACTGGTGATCACCGTAAAGTCATTTTACGCTTCACTGACTTTAACTTGAT GGTACTGGTTATGGTGATTATGTCAAAATATATGATGGATTAGAGGAGAATCCACACAAG CTTTTGCGTGTGTTGACAGCTTTTGATTCTCATGCACCTCTTACAGTTGTTTCTTCTTCT GGACAGATAAGGGTACATTTTGTGCTGATAAAGTGAATGCTGCAAGGGGATTTAATGCT ACTTACCAAGTAGATGGGTTCTGTTTGCCATGGGAAATACCTGTGGAGGTAACCTGGGG TGTTATACTGAGCAGCAGCGTTGTGATGGGTATTGGCATTGCCAAATGGAAGGGATGAA ACCAATGTACCATGTGCCAGAAGGAAGAAATTTCCATGTTCCCGAAATGGTGTCTGTTAT CCTCGTTCTGATCGCTGCAACTACCAGAATCATTGCCCAATGGCTCAGATGAAAAAAC TGCTTTTTTTGCCAACCAGGAAATTTCCATTGTAAAACAATCGTTGTGTGTTTGAAAGT TGGGTGTGTGATTCTCAAGATGACTGTGGTGATGGCAGCGATGAAGAAAATTGCCAGTA ATCGTGCCT	
	ORF Start: at 1	ORF Stop: end of sequence
	SEQ ID NO: 138	463 aa
		MW at 52053.1kD
NOV31h, CG51264-08 Protein Sequence	NGALAEHSENVHISGVSTACGETPEQIRAPSGIITSPGWPEYPAKTNCSWFI RANPGEI ITISFQDFDIQGSRRCLNDWLTIEYKNIESYRACGSTIPPPYISSQDHIWIRF HSDN SRKGFRLAYFSGKSEEPNACDQFRCNGKCIPEAWKCNMDECGDSSDEIICAK EANPP TAAAFQPCAYNQFQCLSRFTKVYTCLPESLKCDGNIDCLDLGDEIDCDVPTC GQWLK FY GTFNSPNYPDFYPPGSNCTWLIDTGDHRKVILRFTDFKLDGTGYGDYVKIYD GLEEN PHK LLRVLTAFDShAPLTVVSSSQIRVHFCAKVNAAARGFNATYQVDGFCLPWEI PCGGN WG CYTEQQRCDGYWHCPNGRDETNTMCQKEEFPSCRNGVCYPRSDRCNYQNHCP NGSDE KN CFFCQPGNFHCKNNRCVFESWVCDSQDDCGDSDEENCPVIVP	
	SEQ ID NO: 139	1389 bp
NOV31i, CG51264-09 DNA Sequence	AATGGTGCTCTTGCAGAACATTCTGAAAATGTGCATATTTCAAGGAGTGTCAACTGCTTGT GGAGAGACTCCAGAGCAAATACGAGACCAAGTGGCATAATCACAAGCCCAGGCTGGCCT TCTGAATATCCTGCAAAAACCAACTGTAGCTGGTTCATAAGGGCAAACCCAGGCGAAATC ATTACTATAAGTTTTCAGGATTTTGATATTCAAGGATCCAGAAGGTGCAATTTGGACTGG TTGACAATAGAAACATACAAGAATATTGAAAGTTACAGAGCTTGTGGTTCACACAATTTCCA CCTCCGTATATCTCTTCACAAGACCACATCTGGATTAGGTTTCATTCCGATGACAACATC TCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGAAATCTGAGGAACCAAATTTGTGCT TGTGATCAGTTTCGTTGTGGTAATGGAAAGTGTATACCAGAAGCCTGGAAATGTAATAAC ATGGATGAATGTGGAGATAGTTCCGATGAAGAGATCTGTGCCAAAGAAGCAAATCCTCCA ACTGCTGCTGCTTTTTCAACCCTGTGCTTACAACCAAGTCCAGTGTTTATCCCGTTTTTACC AAAGTTTACACTTGCTCCCCGAATCTTTAAAATGTGATGGGAACATTGACTGCCTTGAC CTAGGAGATGAGATAGACTGTGATGTGCCAACATGTGGGCAATGGCTAAAATATTTTTAT GGTACTTTTTAATTCTCCCAATTATCCAGACTTTTATCCTCCTGGAAGCAATTGCACCTGG TTAATAGACACTGGTGATCACCGTAAAGTCATTTTACGCTTCACTGACTTTAACTTGAT GGTACTGGTTATGGTGATTATGTCAAAATATATGATGGATTAGAGGAGAATCCACACAAG CTTTTGCGTGTGTTGACAGCTTTTGATTCTCATGCACCTCTTACAGTTGTTTCTTCTTCT GGACAGATAAGGGTACATTTTGTGCTGATAAAGTGAATGCTGCAAGGGGATTTAATGCT ACTTACCAAGTAGATGGGTTCTGTTTGCCATGGGAAATACCTGTGGAGGTAACCTGGGG	

	TGTTATACTGAGCAGCAGCGTTGTGATGGGTATTGGCATTGCCCAAATGGAAGGGATGAA ACCAATTGTACCATGTGCCAGAAGGAAGAAATTTCCATGTTCCCGAAATGGTGTCTGTTAT CCTCGTTCTGATCGCTGCAACTACCAGAATCATTGCCCAAATGGCTCAGATGAAAAAAC TGCTTTTTTTGCCAACCCAGGAAATTTCCATTGTAAAAACAATCGTTGTGTGTTTGAAAGT TGGGTGTGTGATTCTCAAGATGACTGTGGTGATGGCAGCGATGAAGAAAATTGCCAGTA ATCGTGCCT		
	ORF Start: at 1		ORF Stop: end of sequence
	SEQ ID NO: 140	463 aa	MW at 52053.1kD
NOV31i, CG51264-09 Protein Sequence	NGALAEHSENVHISGVSTACGETPEQIRAPSGIITSPGWPSEYPAKTNCSWFIIRANPGEI ITISFQDFDIQGSRRCNLDWLTIIETYKNIESYRACGSTIPPPYISSQDHIWIRFHSDDNI SRKGFRLAYFSGKSEEPNCACDQFRCGNKCIPEAWKCNMDECGDSSDEEICAKEANPP TAAAFQPCAYNQFQCLSRFTKVYTCLPESLKCDGNIDCLDLGDEIDCDVPTCGQWLKYFY GTFNSPNYPDFYPPGSNCTWLIDTGDHRKVILRFTDFKLDGTGYGDYVKIYDGLEENPHK LLRVLTAFDShAPLTVSSSGQIRVHFCADKVNAAARGFNATYQVDGFCPLWEIPCGGNWG CYTEQQRCDGYWHCPNGRDETNTMCQKEEFPCSRNGVCYPRSDRCNYQNHCPNGSDEKN CFFCQPGNFHCKNNRCVFESWVCDSDQDDCGDGSDEENCPVIVP		
	SEQ ID NO: 141	1401 bp	
NOV31j, CG51264-10 DNA Sequence	GGTACCAATGGTGCTCTTGACAGAACATTCTGAAAATGTGCATATTTTCAGGAGTGTCAACT ACTTGTGGAGAGACTCCAGGGCAAATACGAGCACCAGTGGCATAATCACAAGCCCAGGC TGGCCTTCTGAATATCCTGCAAAAATCAACTGTAGCTGGTTCATAAGGGCAAACCCAGGC GAAATCATTACTATAAGTTTTTCAGGATTTTGATATTCAAGGATCCAGAAGGTGCAATTTG GACTGGTTGACAATAGAAACATACAAGAATATTGAAAGTTACAGAGCTTGTGGTTCCACA ATTCCACCTCCGTATATCTCTTCAAGACCACATCTGGATTAGGTTTCATTCCGGATGAC AACATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGAAATCTGAGGAACCAAAT TGTGCTTGTGATCAGTTTCGTTGTGTAATGGAAAGTGTATACCAGAAGCCTGGAATGT AATAACATGGATGAATGTGGAGATAGTTCGGATGAAGAGATCTGTGCCAAAGAGCAAAT CTTCCAACCTGCTGCTGCTTTCAACCCTGTGCTTACAACCAAGTTCAGTGTATTATCCCGT TTTACCAAAGTTTACACTTGCTTCCCGCAATCTTAAAATGTGATGGGAACATTGACTGC CTTGACCTAGGAGATGAGATAGACTGTGATGTGCCAATGTGGGCAATGGCTAAAATAT TTTTATGGTACTTTTAATTCTCCCAATTATCCAGACTTTTATCCTCCTGGAAGCAATTGC ACCTGGTTAATAGACACTGGTGATCACCGTAAAGTCATTTTACGCTTCACTGACTTTAAA CTTGATGGTACTGGTTATGGTGATTATGTCAAAATATATGATGGATTAGAGGAGAATCCA CACAAGCTTTTGCCTGTGTTGACAGCTTTTGATTCTCATGCACCTCTTACAGTTGTTTCT TCTTCTGGACAGATAAGGGTACATTTTGTGCTGATAAAGTGAATGCTGCAAGGGGATTT AATGCTACTTACCAAGTAGATGGGTTCTGTTTGCCATGGGAAATACCCTGTGGAGGTAAC TGGGGGTGTTTACTGAGCAGCAGCGTTGTGATGGGTATTGGCATTGCCCAAATGGAAGG GATGAAACCAATTGTACCATGTGCCAGAAGGAAGAATTTCCATGTTCCCGAAATGGTGTC TGTTATCCTCGTTCTGATCGTGCACCTACCAGAATCATTGCCCAAATGGCTCAGATGAA AAAAACTGCTTTTTTTGCCAACCCAGGAAATTTCCATTGTAAAAACAATCGTTGTGTGTTT GAAAGTTGGGTGTGTGATTCTCAAGATGACTGTGGTGATGGCAGCGATGAAGAAAATTGC CCAGTAATCGTGCCTCGGCCG		
	ORF Start: at 7		ORF Stop: at 1396
	SEQ ID NO: 142	463 aa	MW at 52023.1kD
NOV31j, CG51264-10 Protein Sequence	NGALAEHSENVHISGVSTTCGETPGQIRAPSGIITSPGWPSEYPAKINCSWFIIRANPGEI ITISFQDFDIQGSRRCNLDWLTIIETYKNIESYRACGSTIPPPYISSQDHIWIRFHSDDNI SRKGFRLAYFSGKSEEPNCACDQFRCGNKCIPEAWKCNMDECGDSSDEEICAKEANPP TAAAFQPCAYNQFQCLSRFTKVYTCLPESLKCDGNIDCLDLGDEIDCDVPTCGQWLKYFY GTFNSPNYPDFYPPGSNCTWLIDTGDHRKVILRFTDFKLDGTGYGDYVKIYDGLEENPHK LLRVLTAFDShAPLTVSSSGQIRVHFCADKVNAAARGFNATYQVDGFCPLWEIPCGGNWG CYTEQQRCDGYWHCPNGRDETNTMCQKEEFPCSRNGVCYPRSDRCNYQNHCPNGSDEKN CFFCQPGNFHCKNNRCVFESWVCDSDQDDCGDGSDEENCPVIVP		
	SEQ ID NO: 143	1401 bp	
NOV31k, CG51264-11 DNA Sequence	GGTACCAATGGTGCTCTTGACAGAACATTCTGAAAATGTGCATATTTTCAGGAGTGTCAACT GCTTGTGGAGAGACTCCAGAGCAAATACGAGCACCAGTGGCATAATCACAAGCCCAGGC TGGCCTTCTGAATATCCTGCAAAAACCAACTGTAGCTGGTTCATAAGGGCAAACCCAGGC GAAATCATTACTATAAGTTTTTCAGGATTTTGATATTCAAGGATCCAGAAGGTGCAATTTG		

	GACTGGTTGACAATAGAAACATACAAGAATATTGAAAGTTACAGAGCTTGTGGTTCCACA ATTCCACCTCCGTATATCTCTTCAAGACCACATCTGGATTAGGTTTCATTCCGGATGAC AACATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGAAATCTGAGGAACCAAAT TGTGCTTGTGATCAGTTTCGTTGTGGTAATGGAAGTGTATACCAGAAGCCTGGAAATGT AATAACATGGATGAATGTGGAGATAGTTCGATGAAGAGATCTGTGCCAAAGAAGCAAAT CCTCCAACCTGCTGCTGCTTTTCAACCTGTGCTTACAACCAGTTCAGTGTTTATCCCGT TTTACCAAAGTTTACACTTGCCTCCCCGAATCTTTAAATGTGATGGGAACATTGACTGC CTTGACCTAGGAGATGAGATAGACTGTGATGTGCCAACATGTGGGCAATGGCTAAAATAT TTTTATGGTACTTTTAATTCTCCAATTATCCAGACTTTTATCCTCCTGGAAGCAATTGC ACCTGGTTAATAGACACTGGTGATCACCGTAAAGTCATTTTACGCTTCACTGACTTTAAA CTTGATGGTACTGGTTATGGTGATTATGTCAAAATATATGATGGATTAGAGGAGAATCCA CACAAGCTTTTGCCTGTGTTGACAGCTTTTGATTCTCATGCACCTCTTACAGTTGTTTCT TCTTCTGGACAGATAAGGGTACATTTTTGTGCTGATAAAGTGAATGCTGCAAGGGGATTT AATGCTACTTACCAAGTAGATGGGTTCTGTTTGCCATGGGAAATACCCTGTGGAGGTAAC TGGGGGTGTTATACTGAGCAGCAGCGTTGTGATGGGTATTGGCATTGCCCAAATGGAAGG GATGAAACCAATTGTACCATGTGCCAGAAGGAAGAATTTCCATGTTCCCGAAATGGTGTC TGTTATCCTCGTTCTGATCGCTGCAACTACCAGAATCATTGCCCAAATGGCTCAGATGAA AAAAACTGCTTTTTTTGCCAACCAGGAAATTTCCATTGTAAAAACAATCGTTGTGTGTTT GAAAGTTGGGTGTGTGATTCTCAAGATGACTGTGGTGATGGCAGCGATGAAGAAAATTGC CCAGTAATCGTGCCTCGGCCG		
	ORF Start: at 7		ORF Stop: at 1396
	SEQ ID NO: 144	463 aa	MW at 52053.1kD
NOV31k, CG51264-11 Protein Sequence	NGALAEHSENVHISGVSTACGETPEQIRAPSGIITS PGWPSEYPAKTNCSWFI RANPGEI ITISFQDFDIQGSRRCNLDWLT IETKYNI ESYRACGSTIPPPYISSQDHIWIRFHSSDNI SRKGFR LAYFSGKSEEPNCADQFR CGNGKCIPEAWKCNMDECGDSSDEI CAKEANPP TAAAFQPCAYNQFQLSRFTKVYTCLPESLKCDGNIDCLDLGDEIDCDVPTCGQWLKYFY GTFNSPNYPDFYPPGSNCTWLIDTGDHRKVILRFTDFKLDGTGYGDYVKIYDGLEENPHK LLRVLTAFD SHAPLTVVSSSGQIRVHF CADKVNAARGFNATYQVDGFCLPWEIPCGGNWG CYTEQQRCDGYWHCPNGRDETNTMCQKEFP CSRNGVCYPRSDRCNYQNHC PNGSDEKN CFFCQPGNFHCKNNRCVFESWVCDSDDCGDGSDEENCPVIVP		
	SEQ ID NO: 145	1401 bp	
NOV31I, CG51264-12 DNA Sequence	GGTACCAATGGTGCTCTTG CAGAACATTCTGAAAATGTGCATATTT CAGGAGTGTCAACT GCTTGTGGAGAGACTCCAGAGCAAATACGAGCACCAAGTGGCATAATCACAAGCCAGGC TGGCCTTCTGAATATCCTGCAAAAATCAACTGTAGCTGGTTCATAAGGCAAAACCAGGC GAAATCATTACTATAAGTTTT CAGGATTTTGATATTCAAGGATCCAGAAGGTGCAATTTG GACTGGTTGACAATAGAAACATACAAGAATATTGAAAGTTACAGAGCTTGTGGTTCCACA ATTCCACCTCCGTATATCTCTTCAAGACCACATCTGGATTAGGTTTCATTCCGGATGAC AACATCTCTAGAAAGGGTTTCAGACTGGCATATTTT CAGGGAAATCTGAGGAACCAAAT TGTGCTTGTGATCAGTTTCGTTGTGGTAATGGAAGTGTATACCAGAAGCCTGGAAATGT AATAACATGGATGAATGTGGAGATAGTTCGATGAAGAGATCTGTGCCAAAGAAGCAAAT CCTCCAACCTGCTGCTGCTTTTCAACCTGTGCTTACAACCAGTTCAGTGTTTATCCCGT TTTACCAAAGTTTACACTTGCCTCCCCGAATCTTTAAATGTGATGGGAACATTGACTGC CTTGACCTAGGAGATGAGATAGACTGTGATGTGCCAACATGTGGGCAATGGCTAAAATAT TTTTATGGTACTTTTAATTCTCCCAATTATCCAGACTTTTATCCTCCTGGAAGCAATTGC ACCTGGTTAATAGACACTGGTGATCACCGTAAAGTCATTTTACGCTTCACTGACTTTAAA CTTGATGGTACTGGTTATGGTGATTATGTCAAAATATATGATGGATTAGAGGAGAATCCA CACAAGCTTTTGCCTGTGTTGACAGCTTTTGATTCTCATGCACCTCTTACAGTTGTTTCT TCTTCTGGACAGATAAGGGTACATTTTGTGCTGATAAAGTGAATGCTGCAAGGGGATTT AATGCTACTTACCAAGTAGATGGGTTCTGTTTGCCATGGGAAATACCCTGTGGAGGTAAC TGGGGGTGTTATACTGAGCAGCAGCGTTGTGATGGGTATTGGCATTGCCCAAATGGAAGG GATGAAACCAATTGTACCATGTGCCAGAAGGAAGAATTTCCATGTTCCCGAAATGGTGTC TGTTATCCTCGTTCTGATCGCTGCAACTACCAGAATCATTGCCCAAATGGCTCAGATGAA AAAAACTGCTTTTTTTGCCAACCAGGAAATTTCCATTGTAAAAACAATCGTTGTGTGTTT GAAAGTTGGGTGTGTGATTCTCAAGATGACTGTGGTGATGGCAGCGATGAAGAAAATTGC CCAGTAATCGTGCCTCGGCCG		
	ORF Start: at 7		ORF Stop: at 1396
	SEQ ID NO: 146	463 aa	MW at 52065.2kD

NOV31l, CG51264-12 Protein Sequence	NGALAEHSENVHISGVSTACGETPEQIRAPSGIITSPGWPSEYPAKINCSWFIRANPGEI ITISFQDFDIQGSRRCNLDWLTIIETYKNIESYRACGSTIPPPYISSQDHIWIRFHSDDNI SRKGFRLAYFSGKSEEPNCADQFRCGNKCIPEAWKCNMDECGDSSDEEICAKEANPP TAAAFQPCAYNQFQCLSRFTKVYTCLPESLKCDGNIDCLDLGDEIDCDVPTCGQWLKYFY GTFNSPNYPDFYPPGSNCTWLIDTGDHRKVIILRFTDFKLDGTGYGDYVKIYDGLEENPHK LLRVLTAFDASHAPLTVVSSSGQIRVHFCADKVNAARGFNATYQVDGFCLPWEIPCNGNWG CYTEQQRCDGYWHCPNGRDETNTMCQKEEPCSRNGVCYPRSDRCNYQNHCPNGSDEKN CFFCQPGNFHCKNNRCVFESWVCDSDQDDCGDGSDEENCPVIVP		
	SEQ ID NO: 147	1401 bp	
NOV31m, CG51264-13 DNA Sequence	GGTACCAATGGTGCTCTTGCAGAACATTCTGAAAATGTGCATATTTTCAGGAGTGTCAAC TGCTTGTGGAGAGACTCCAGAGCAAATACGAGCACCAAGTGGCATAATCACAAGCCCAG GCTGGCCTTCTGAATATCCTGCAAAAATCAACTGTAGCTGGTTCAATAGGGCAAACCCA GGCGAAATCATTACTATAAGTTTTCAGGATTTTGATATTCAAGGATCCAGAAGGTGCAA TTTGGACTGGTTGACAATAGAAACATACAAGAATATTGAAAGTTACAGAGCTTGTGGTT CTACAATTCACCTCCGTATATCTCTTACAAGACCACATCTGGATTAGGTTTCATTCTG GATGACAACATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGAAATCTGAGGA ACCAAATTGTGCTTGTGATCAGTTTCGTTGTGGTAATGGAAAGTGTATACCAGAAGCCT GGAAATGTAATAACATGGATGAATGTGGAGATAGTTCGGATGAAGAGATCTGTGCCAAA GAAGCAAATCCTCCAAGTGTCTGCTTTTCAACCCTGTGCTTACAAACCAAGTTCAGTG TTTATCCCGTTTACCAAAGTTTACACTTGCTCCCGAATCTTTAAATGTGATGGGA ACATTGACTGCCTTGACCTAGGAGATGAGATAGACTGTGATGTGCCAATCTGTGGCAA TGGCTAAAATATTTTATGGTACTTTTAATTCTCCCAATTATCCAGACTTTTATCCTCC TGAAGCAATTGCACCTGGTTAATAGACACTGGTGATCACCGTAAAGTCATTTTACGCT TCACTGACTTTAACTTGATGGTACTGGTTATGGTGATTATGTCAAATATATGATGGA TTAGAGGAGAATCCACACAAGCTTTTGCCTGTGTTGACAGCTTTTGATTCTCACGACC TCTTACAGTTGTTTCTTCTCTGGACAGATAAGGGTACATTTTGTGCTGATAAAGTGA ATGCTGCAAGGGGATTTAATGCTACTTACCAAGTAGATGGGTTCTGTTGCCATGGGAA ATACCCTGTGGAGGTAAGTGGGGTGTATACTGAGCAGCAGCGTTGTGATGGGTATTG GCATTGCCCAAATGGAAGGGATGAAACCAATTGTACCATGTGCCAGAAGGAAGAAATTC CATGTTCCCGAAATGGTGTCTGTTATCCTCGTTCTGATCGCTGCACTACCAAGATCAT TGCCCAAATGGCTCAGATGAAAAAACTGCTTTTTTGGCAACCAGGAAATTTCCATTG TAAAAACAATCGTTGTGTGTTTGAAGTTGGGTGTGTGATTCTCAAGATGACTGTGGTG ATGGCAGCGATGAAGAAAATTGCCAGTAATCGTGCTCGGCCG		
	ORF Start: at 7		ORF Stop: at 1396
	SEQ ID NO: 148	463 aa	MW at 52065.2kD
NOV31m, CG51264-13 Protein Sequence	NGALAEHSENVHISGVSTACGETPEQIRAPSGIITSPGWPSEYPAKINCSWFIRANPGEI IITISFQDFDIQGSRRCNLDWLTIIETYKNIESYRACGSTIPPPYISSQDHIWIRFHSDD NISRKGFRLAYFSGKSEEPNCADQFRCGNKCIPEAWKCNMDECGDSSDEEICAKEANPP TAAAFQPCAYNQFQCLSRFTKVYTCLPESLKCDGNIDCLDLGDEIDCDVPTCGQWL KYFYGTFNSPNYPDFYPPGSNCTWLIDTGDHRKVIILRFTDFKLDGTGYGDYVKIYDGLE ENPHKLLRVLTAFDASHAPLTVVSSSGQIRVHFCADKVNAARGFNATYQVDGFCLPWEIP CGGNWGCYTEQQRCDGYWHCPNGRDETNTMCQKEEPCSRNGVCYPRSDRCNYQNHCP NGSDEKNCFQCQPGNFHCKNNRCVFESWVCDSDQDDCGDGSDEENCPVIVP		
	SEQ ID NO: 149	1401 bp	
NOV31n, CG51264-14 DNA Sequence	GGTACCAATGGTGCTCTTGCAGAACATTCTGAAAATGTGCATATTTTCAGGAGTGTCAACT GCTTGTGGAGAGACTCCAGAGCAAATACGAGCACCAAGTGGCATAATCACAAGCCCAGGC TGGCCTTCTGAATATCCTGCAAAAATCAACTGTAGCTGGTTCAATAGGGCAAACCCAGGC GAAATCATTACTATAAGTTTTCAGGATTTTGATATTCAAGGATCCAGAAGGTGCAATTTG GACTGGTTGACAATAGAAACATACAAGAATATTGAAAGTTACAGAGCTTGTGGTTCCACA ATTCCACCTCCGTATATCTCTTACAAGACCACATCTGGATTAGGTTTCATTCCGATGAC AACATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGGGAAATCTGAGGAACCAAT TGTGCTTGTGATCAGTTTCGTTGTGGTAATGGAAAGTGTATACCAGAAGCTGGAAATGT AATAACATGGATGAATGTGGAGATAGTTCGGATGAAGAGATCTGTGCCAAAGAAGCAAAT CCTCCAAGTGTGCTGCTTTTCAACCCTGTGCTTACAACCAGTTCAGTGTATATCCCGT TTACCAAAGTTTACACTTGCTCCCGAATCTTTAAATGTGATGGGAACATTGACTGC CTTGACCCAGGAGATGAGATAGACTGTGATGTGCCAATGTGGGAATGGCTAAAATAT TTTTATGGTACTTTTAATTCTCCCAATTATCCAGACTTTTATCCTCCTGAAGCAATTGC		

	ACCTGGTTAATAGACACTGGTGATACCGTAAAGTCATTTTACGCTTCACTGACTTTAAACCTGATGGTACTGGTTATGGTGATTATGTCAAATATATGATGGATTAGAGGAGAATCCACACAAGCTTTTGC		
	ORF Start: at 7		
	ORF Stop: at 1396		
	SEQ ID NO: 150	463 aa	MW at 52050.1kD
NOV31n, CG51264-14 Protein Sequence	NGALAEHSENVHISGVSTACGETPEQIRAPSGIITS		
	SEQ ID NO: 151		
	2592 bp		
NOV31o, CG51264-15 DNA Sequence	GGCCTGTCGCTGGAGCACAAAGAGTCTCCGCGGTGGAGGTCTGCGTTGCTCTTGCTTTT		

	AGCACGTCACCAGCTTACAAGTGCCTCAGTCGTATGACTCAGGGGCTACGCTGGGTACG TTTTACATTAGGACGATCAAGTTCCCTAAGTCAGAACCAGAGTCCTTTGAGACAACTTGA TAATGGGGTAAGTGAAGAGAAGATGATGATGATGTTGAAATGCTAATTCGAATTTCTGA TGGATCTTCAGACTTTGATGTGAATGACTGCTCCAGACCTCCTCTTGATCTTGCCCTCAGA TCAAGGACAAGGGCTTAGACAACCATATAATGCAACAAATCCTGGAGTAAGGCCAAGTAA TCGAGATGGCCCCCTGTGAGCGCTGTGGTATTGTCCACACTGCCAGATACCCAGACTGT CTTAGAAGTAACACTGAAAAACGAAACGAGTGGTGATGAGGCTTTGTTACTTTGTTAGGT ACGAATCACATA		
	ORF Start: at 2		ORF Stop: TAG at 2576
	SEQ ID NO: 152	858 aa	MW at 94777.5kD
NOV31o, CG51264-15 Protein Sequence	ACRWSTKESPRWRSALLLLFLAGVYNGALAEHSENVHISGVSTACGETPEQIRAPSGII TSPGWPSEYPAKINCSWFIRANPGEIITISFQDFDIQGSRRCNLDWLTITYKNIESYRA CGSTIPPPYISSQDHIWIRFHSDDNISRKGFRLAYFSGKSEEPNCACDQFCGNGKCIPE AWKCNMDECGDSSDEEICAKEANPPTAAAFQPCAYNQFQCLSRFTKVYTCLPESLKCDG NIDCLDLGDEIDCDVPTCGQWLKYFYGTFFNSPNYPDFYPPGSNCTWLIDTGDHRKVI LRF TDFKLDGTGYGDYVKIYDGLEENPHKLLRLVLTAFDSHAPLTVVSSSGQIRVHFCADKVN ARGFNATYQVDGFCLPWEIPCGGNWGCYTEQQRCDGYWHCPNGRDETNTCMQKEEFPCS RNGVCYPRSDRCNYQNHCPNGSDEKNCFFCQPGNFHCKNNRCVFESWVCDSQDDCGDGS EENCPIVIVPTRVITA AVIGSLICGLLLVIALGCTCKLYSLRMFERRSFETQLSRVEAELL RREAPPSYQLIAQGLIPPVEDFPVCSNPQASVLENLRLAVRSQLGFTSVRLPMAGRSSN IWNRI FNFARSRHSGSLALVSADGDEVVPSQSTSREPERNHTRSLFVSDEDDTD TENER RDMAGASGGVAAPLPQKVPPTTAVEATVGACASSSTQSTRGGHADNDRDVSVEPPSVSP ARHQLTSALSRMTQGLRWVFTLGRSSSLSONQSPLRQLDNGVSGREDDDDVEMLIPI SD GSSDFD VNDCSRPPDLASDQGGQLRQYPNATNPGVRPSNRDGPCERCIGIVHTAQIPDTC LEVTLKNETSGDEALLC		
	SEQ ID NO: 153	2560 bp	
NOV31p, CG51264-16 DNA Sequence	TATGGCCTGTCGCTGGAGCACAAAAGAGTCTCCGCGGTGGAGGTCTGCGTTGCTCTTGCT TTTCCTCGCTGGGGTGACGCTTGTTGGAGAGACTCCAGGGCAAATACGAGCACCAGTGG CATAATCACAAGCCAGGCTGGCCTTCTGAATATCCTGCAAAAATCAACTGTAGCTGGTT CATAAGGGCAAACCCAGGCGAAATCATTACTATAAGTTTTTCAGGATTTTGATATTCAAGG ATCCAGAAGGTGCAATTTGGACTGGTTGACAATAGAAACATACAAGAATATTGAAAGTTA CAGAGCTTTGTTGTTCCACAATTCACCTCCGTATATCTCTTCAAGACCACATCTGGAT TAGGTTTCATTCCGATGACAACATCTCTAGAAAGGGTTTCAGACTGGCATATTTTTCAGG GAAATCTGAGGAACCAAATTTGTGCTTGTGATCAGTTTCGTTGTGGTAATGGAAAGTGAT ACCAGAAGCCTGGAATGTAATAACATGGATGAATGTGGAGATAGTTCGATGAAGAGAT CTGTGCCAAAGAAGCAAATCTCCAACCTGCTGCTGCTTTTCAACCCTGTGCTTTACAACCA GTTCCAGTGTTTATCCGTTTACCAAAGTTTACACTTGCTCCCGCAATCTTTAAAATG TGATGGGAACATTGACTGCCTTGACCTAGGAGATGAGATAGACTGTGATGTGCCAATG TGGGCAATGGCTAAAATATTTTATGGTACTTTTAATTCTCCCAATTATCCAGACTTTTA TCCTCTCGGAAGCAATTGCACCTGGTTAATAGACACTGGTGATCACCCTAAAGTCATTTT ACGCTTCACTGACTTTAACTTGATGGTACTGGTTATGGTGATTATGTCAAATATATGA TGGATTAGAGGAGAATCCACAAGCTTTTGCGTGTGTTGACAGCTTTTGATTCTCATGC ACCTCTTACAGTTGTTTCTTCTTCTGGACAGATAAGGGTACATTTTGTGCTGATAAAGT GAATGCTGCAAGGGGATTTAATGCTACTTACCAAGTAGATGGGTTCTGTTTGCCATGGGA AATACCCTGTGGAGGTAACGGGGGTGTTATACTGAGCAGCAGCGTTGTGATGGGTATTG GCATTGCCCAAATGGAAGGGATGAAACCAATTGTACCATGTGCCAGAAGGAAGAATTTCC ATGTTCCCGAAATGGTGTCTGTTATCCTCGTTCTGATCGCTGCAACTACCAGAATCATTG CCCAAATGGCTCAGATGAAAAAACTGCTTTTTTGGCAACCAGGAAATTTCCATTGTAA AAACAATCGTTGTGTGTTTGAAGTTGGGTGTGTGATTCTCAAGATGACTGTGGTGATGG CAGCGATGAAGAAAATTGCCCAGTAATCGTGCCTACAAGAGTCATCACTGCTGCCGTCAT AGGGAGCCTCATCTGTGGCCTGTTACTCGTCATAGCATTGGGATGTACTGTAAAGCTTTA TTCTCTGAGAATGTTTGAAGAAGATCATTTGAAACACAGTTGTCAAGATGGGAAGCAGA ATTGTTAAGAAGAGAAGCTCCTCCCTCGTATGGACAATTGATTGCTCAGGGTTTAATTCC ACCAGTTGAAGATTTTCTGTTTGTTCACCTAATCAGGCTTCTGTTTGGAAAACTGAG GCTAGCGGTACGATCTCAGCTTGGATTACTTCAGTCAGGCTTCTATGGCAGGCAGATC AAGCAACATTTGGAACCGTATTTTAAATTTTGAAGATCACGTCATTCTGGGTCAATTGGC TTTGGTCTCAGCAGATGGAGATGAGGTGTCCCTAGTCAGATCAGGTAGAGAACCTGA GAGAAATCATACTACAGAAGTTTGTTTCCGTTGGAGTCTGATGATACAGACAGAAAA		

	TGAGAGAAGAGATATGGCAGGAGCATCTGGTGGGGTTGCAGCTCCTTTGCCTCAAAAAGT CCCTCCCAACGGCAGTAGAAGCGACAGTAGGAGCATGTGCAAGTTCTCAACTCAGAG TACCCGAGGTGGTCATGCAGATAATGGAAGGGATGTGACAAGTGTGGAACCCCAAGTGT GAGTCCAGCACGTACCAGCTTACAAGTGCACACTCAGTCGTATGACTCAGGGGCTACGCTG GGTACGTTTTACATTAGGACGATCAAGTTCCTAAGTCAGAACCAGAGTCTTTGAGACA ACTTGATAATGGGGTAAGTGGAAAGAGAAGATGATGATGATGTTGAAATGCTAATTCCAAT TTCTGATGGATCTTCAGACTTTGATGTGAATGACTGCTCCAGACCTCTTCTTGATCTTGC CTCAGATCAAGGACAAGGGCTTAGACAACCATATAATGCAACAAATCCTGGAGTAAGGCC AAGTAATCGAGATGGCCCTGTGAGCGCTGTGGTATTGTCCACACTGCCCAGATACCAGA CACTTGCTTAGAAGTAACACTGAAAAACGAAACGAGTGATGATGAGGCTTTGTTACTTTG TTAGGTACGAATCACATAAGGGCGATTCCAGCACCTGGCT		
	ORF Start: ATG at 2		ORF Stop: TAG at 2522
	SEQ ID NO: 154	840 aa	MW at 93049.7kD
NOV31p, CG51264-16 Protein Sequence	MACRWSTKESPRWRSALLLFLAGVYACGETPGQIRAPSGIITSPGWSEYPKINCSWF IRANPGEIITISFQDFDIQSSRCNLDWLTITYKNIESYRACGSTIPPPYISSQDHIWI RFHSDDNISRKGFRLAYFSGKSEEPNCADQFRCGNGKCIPEAWKCNNMDECGDSSDEEI CAKEANPPTAAAFQPCAYNQFQCLSRFTKVYTCLPESLKCDGNIDCLDLGDEIDCDVPTC GQWLKIFYGTFNPNYPDFYPPGSNCTWLIDTGDHRKVILRFTDFKLDGTGYGDYVKIYD GLEENPHKLLRVLTAFDSHAPLTVVSSSGQIRVHFCADKVNAAARGFNATYQVDGFCLPWE IPCGGNWGCYTEQQRCDCGYWHCPNGRDETNCTMCQKEEPCSRNGVCYPRSDRCNYQNH PNGSDEKNCFPCQPGNFHCKNNRCVFESWVCDSDQDCGDSDEENCPIVIVPTRVITA AVIGSLICGLLLVIALGCTCKLYSLRMFERRSFETQLSRVEAELLRREAPPSYGQLIAQGLIP PVEDFPVCSNPQASVLENLRLAVRSQLGFTSVRLPMAGRSSNIWNRI FNFARSRHSGSLA LVSADGDEVVPSQSTSREPERNHTRSLFSVESDDTDENERRDMAGASGGVAAPLPQKV PPTTAVEATVGACASSSTQSTRGGHADNGRDVTSVEPPSVSPARHQLTSALSRMTQGLRW VRFTLGRSSSLSONQSPLRQLDNGVSGREDDDDVEMLIPISDGSDFDVNDCSRPLDLA SDQGQGLRQPYNATNPGVRPSNRDGPCERCGIVHTAQIPDTCLEVTLKNETSDDEALLC		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 31B.

Table 31B. Comparison of NOV31a against NOV31b through NOV31p.		
Protein Sequence	NOV31a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV31b	1..423 1..442	422/442 (95%) 422/442 (95%)
NOV31c	1..423 1..423	422/423 (99%) 422/423 (99%)
NOV31d	1..840 1..840	826/840 (98%) 826/840 (98%)
NOV31e	1..840 1..837	815/840 (97%) 816/840 (97%)
NOV31f	1..423 40..462	422/423 (99%) 422/423 (99%)
NOV31g	1..840 1..859	826/859 (96%) 826/859 (96%)
NOV31h	27..471 19..463	430/445 (96%) 430/445 (96%)

NOV31i	27..471 19..463	430/445 (96%) 430/445 (96%)
NOV31j	28..471 20..463	429/444 (96%) 429/444 (96%)
NOV31k	27..471 19..463	430/445 (96%) 430/445 (96%)
NOV31l	27..471 19..463	431/445 (96%) 431/445 (96%)
NOV31m	27..471 19..463	431/445 (96%) 431/445 (96%)
NOV31n	27..471 19..463	429/445 (96%) 430/445 (96%)
NOV31o	2..840 1..858	823/858 (95%) 823/858 (95%)
NOV31p	1..840 1..840	825/840 (98%) 825/840 (98%)

Further analysis of the NOV31a protein yielded the following properties shown in Table 31C.

Table 31C. Protein Sequence Properties NOV31a	
PSort analysis:	0.4600 probability located in plasma membrane; 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in outside
SignalP analysis:	Cleavage site between residues 28 and 29

A search of the NOV31a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 31D.

Table 31D. Geneseq Results for NOV31a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV31a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB70544	Human PRO14 protein sequence SEQ ID NO:28 - <i>Homo sapiens</i> , 840 aa. [WO200110902-A2, 15-FEB-2001]	1..840 1..840	840/840 (100%) 840/840 (100%)	0.0

AAO20441	Protein of the human cancer suppressor gene 98 - <i>Homo sapiens</i> , 894 aa. [CN1328030-A, 26-DEC-2001]	1..840 36..894	840/859 (97%) 840/859 (97%)	0.0
AAU14316	Human novel protein #187 - <i>Homo sapiens</i> , 859 aa. [WO200155437-A2, 02-AUG-2001]	1..840 1..859	840/859 (97%) 840/859 (97%)	0.0
AAB42317	Human ORFX ORF2081 polypeptide sequence SEQ ID NO:4162 - <i>Homo sapiens</i> , 859 aa. [WO200058473-A2, 05-OCT-2000]	1..840 1..859	840/859 (97%) 840/859 (97%)	0.0
AAV02381	Polypeptide identified by the signal sequence trap method - <i>Homo sapiens</i> , 859 aa. [WO9918126-A1, 15-APR-1999]	1..840 1..859	840/859 (97%) 840/859 (97%)	0.0

In a BLAST search of public sequence databases, the NOV31a protein was found to have homology to the proteins shown in the BLASTP data in Table 31E.

Table 31E. Public BLASTP Results for NOV31a				
Protein Accession Number	Protein/Organism/Length	NOV31a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAC33422	Sequence 27 from Patent WO0110902 - <i>Homo sapiens</i> (Human), 840 aa.	1..840 1..840	840/840 (100%) 840/840 (100%)	0.0
Q9Y561	ST7 protein - <i>Homo sapiens</i> (Human), 859 aa.	1..840 1..859	840/859 (97%) 840/859 (97%)	0.0
Q9BE74	Hypothetical 73.8 kDa protein - <i>Macaca fascicularis</i> (Crab eating macaque) (<i>Cynomolgus</i> monkey), 672 aa.	169..840 1..672	663/672 (98%) 666/672 (98%)	0.0
CAC38967	Sequence 19 from Patent WO0119856 - <i>Homo sapiens</i> (Human), 430 aa.	1..423 1..423	422/423 (99%) 422/423 (99%)	0.0
CAC33423	Sequence 29 from Patent WO0110902 - <i>Homo sapiens</i> (Human), 449 aa.	1..423 1..442	422/442 (95%) 422/442 (95%)	0.0

PFam analysis predicts that the NOV31a protein contains the domains shown in Table 31F.

Table 31F. Domain Analysis of NOV31a			
Pfam Domain	NOV31a Match Region	Identities/ Similarities for the Matched Region	Expect Value
CUB	28..137	41/119 (34%) 89/119 (75%)	3.9e-31
ldl_recept_a	145..183	19/43 (44%) 30/43 (70%)	2.1e-10
ldl_recept_a	194..237	17/47 (36%) 27/47 (57%)	6.6e-05
CUB	240..350	42/120 (35%) 83/120 (69%)	6.6e-23
ldl_recept_a	354..393	15/43 (35%) 23/43 (53%)	0.072
ldl_recept_a	394..431	17/44 (39%) 27/44 (61%)	0.045
ldl_recept_a	432..468	21/43 (49%) 32/43 (74%)	1.4e-11

Example 32.

- 5 The NOV32 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 32A.

Table 32A. NOV32 Sequence Analysis			
	SEQ ID NO: 155	2365 bp	
NOV32a, CG52423-01 DNA Sequence	ACGCGTTGATATCCGCCCGAGCTCCGGCGCAGCTCCTCCACCTTGGAGCTCATGAGAG CAGGCCTGGTGGTGAGCAGGACGGTGCACCGACGGCGGGATCGAGCAATGGGTCTGG CCATGGAGCACGGAGGGTCTACGCTCGGGCGGGGGGAGCTCTCGGGGCTGCTGGTATT ACCTGCGCTACTTCTTCTCTTCTGCTCTCCCTCATCCAATTCTCATCATCTGGGGCTCG TGCTCTTCATGGTCTATGGCAACGTGCACGTGAGCACAGAGTCCAACCTGCAGGCCACCG AGCGCCGAGCCGAGGGCCTATACAGTCAGCTCCTAGGGCTCACGGCCTCCAGTCCAACCT TGACCAAGGAGCTCAACTTCACCACCGCGCCAAGGATGCCATCATGCAGATGTGGCTGA ATGCTCGCCGCGACCTGGACCGCATCAATGCCAGCTTCCGCCAGTGCCAGGGTGACCGGG TCATCTACACGAACAATCAGAGGTACATGGCTGCCATCATCTTGAGTGAGAAGCAATGCA GAGATCAATTCAAGGACATGAACAAGAGCTGCGATGCCTTGCTCTTCATGCTGAATCAGA AGGTGAAGACGCTGGAGGTGGAGATAGCCAAGGAGAAGACCATTTGACTAAGGATAAGG AAAGCGTGCTGCTGAACAAACGCGTGGCGGAGGAACAGCTGGTTGAATGCGTGAAAACCC GGGAGCTGCAGCACCAAGAGCGCCAGCTGGCCAAGGAGCAACTGCAAAAAGGTGCAAGCCC TCTGCCTGCCCTGGACAAGGACAAGTTTGAGATGGACCTTCGTAACCTGTGGAGGGACT CCATTATCCACGCAGCCTGGACAACCTGGGTTACAACCTTACCATCCCTGGGCTCGG AATTGGCCTCCATCCGCAGAGCCTGCGACCACATGCCAGCCTCATGAGCTCCAAGGTGG		

	AGGAGCTGGCCCGGAGCCTCCGGGCGGATATCGAACGCGTGCCCGCGAGAACTCAGACC TCCAACGCCAGAAGCTGGAAGCCCAGCAGGGCCTGCGGGCCAGTCAGGAGGCGAAACAGA AGGTGGAGAAGGAGGCTCAGGCCCGGGAGGCCAAGCTCCAAGCTGAATGCTCCCGGCAGA CCAGCTAGCGCTGGAGGAGAAGGCGGTGCTGCGGAAGGAACGAGACAACCTGGCCAAGG AGCTGGAAGAGAAGAAGAGGGAGGCGGAGCAGCTCAGGATGGAGCTGGCCATCAGAACT CAGCCCTGGACACCTGCATCAAGACCAAGTCGCAGCCGATGATGCCAGTGTCAAGGCCCA TGGGCCCTGTCCCAACCCCCAGCCCATCGACCCAGCTAGCCTGGAGGAGTTCAAGAGGA AGATCCTGGAGTCCCAGAGGCCCTGTCAGGCATCCCTGTAGCCCATCCAGTGGCTGAG GAGGCTCCAGGCCTGAGGACCAAGGGATGGCCGACTCGGCGGTTTGGCGGAGGATGCAGG GATATGCTCACAGCGCCCGACACAACCCCTCCCGCCGCCCCAACCCAGGGGCCACC ATCAGACAACTCCCTGCATGCAAACCCCTAGTACCCTCTCACACCCGCACCCGCGCCTCA CGATCCCTCACCAGAGCACACGGCCGCGGAGATGACGTCACCAAGCAACGGCGCTGAC GTCACATATCACCGTGGTGATGGCGTCACGTGGCCATGTAGACGTCACGAAGAGATATAG CGATGGCGTCGTGCAGATGCAGCAGTCGCACACAGACATGGGGAACCTGGCATGACGTC ACACCGAGATGCAGCAACGACGTCACGGGCCATGTCGACGTCACACATATTAATGTCACA CAGACGCGCGATGGCATCACACAGACGGTGATGATGTCACACACAGACACAGTGACAAC ACACACCATGACAACGACACCTATAGATATGGCACCAACATCACATGCACGCATGCCCTT TCACACACACTTTCTACCCAATTCTCACCTAGTGTACGTTCCCCGACCTTGGCACACG GGCCAAGGTACCCACAGGATCCCATCCCCCTCCCGCACAGCCCTGGGCCCCAGCACCTCCC CTCCTCCAGCTTCCTGGCCTCCAGCCACTTCCTCACCCCCAGTGCCTGGACCCGGAGGT GAGAACAGGAAGCCATTACCTCCGCTCCTTGAGCGTGAGTGTTCCAGGACCCCTCGG GGCCCTGAGCCGGGGGTGAGGGTCACCTGTTGTGCGGAGGGGAGCCACTCCTTCTCCCC AACTCCAGCCCTGCCTGTGGCCGTTGAAATGTTGGTGGCACTTAATAAATATTAGTAA ATCCTTAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 54		ORF Stop: TGA at 1437
	SEQ ID NO: 156	461 aa	MW at 52503.8kD
NOV32a, CG52423-01 Protein Sequence	MRAGLVVSRDGPDDGGIEQMGLAMEHGGSYARAGGSSRGCWYYLRYFFLFVSLIQFLIIL GLVLFMVYGNVHVSTESNLQATERRAEGLYSQLLGLTASQSNLTKELNFTTRAKDAIMQM WLNARRDLDRINASFRCQCGDRVIYTNQRYMAAIILSEKQCRDQFKDMNKSCDALLFML NQKVKTLVEVIAKEKTICTKDKESVLLNKRVAEEQLVECVKTRQLQHQRQLAKEQLQKV QALCLPLDKDFEMDLRNLWRDSIIIPRSLDNLGYNLYHPLGSELASIRRACDHMPSLMSS KVEELARSLRADIERVARENSDLQRQKLEAQQLRASQEAQKVEKEAQAREAKLQAECS RQTQLALEEKAVLRKERDNLAKELEKKRAEQLRMELAIRNSALDTCIKTSQPMMPVS RPMGPVNPQPIDPASLEEFKRKILESQRPPAGIPVAPSSG		

Twenty polymorphic variants of NOV32a have been identified and are shown in Table 41L. Further analysis of the NOV32a protein yielded the following properties shown in Table 32B.

Table 32B. Protein Sequence Properties NOV32a	
PSort analysis:	0.7900 probability located in plasma membrane; 0.6000 probability located in nucleus; 0.3000 probability located in microbody (peroxisome); 0.3000 probability located in Golgi body
SignalP analysis:	Cleavage site between residues 70 and 71

- A search of the NOV32a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 32C.

Table 32C. Geneseq Results for NOV32a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV32a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB42154	Human ORFX ORF1918 polypeptide sequence SEQ ID NO:3836 - <i>Homo sapiens</i> , 479 aa. [WO200058473-A2, 05-OCT-2000]	1..461 19..479	461/461 (100%) 461/461 (100%)	0.0
AAM41619	Human polypeptide SEQ ID NO 6550 - <i>Homo sapiens</i> , 457 aa. [WO200153312-A1, 26-JUL-2001]	7..461 3..457	454/455 (99%) 454/455 (99%)	0.0
AAE06600	Human protein having hydrophobic domain, HP10787 - <i>Homo sapiens</i> , 442 aa. [WO200149728-A2, 12-JUL-2001]	20..461 1..442	442/442 (100%) 442/442 (100%)	0.0
AAM39833	Human polypeptide SEQ ID NO 2978 - <i>Homo sapiens</i> , 442 aa. [WO200153312-A1, 26-JUL-2001]	20..461 1..442	439/442 (99%) 439/442 (99%)	0.0
AA Y12280	Human 5' EST secreted protein SEQ ID NO:311 - <i>Homo sapiens</i> , 105 aa. [WO9906548-A2, 11-FEB- 1999]	20..124 1..105	104/105 (99%) 104/105 (99%)	3e-54

In a BLAST search of public sequence databases, the NOV32a protein was found to have homology to the proteins shown in the BLASTP data in Table 32D.

Table 32D. Public BLASTP Results for NOV32a				
Protein Accession Number	Protein/Organism/Length	NOV32a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAD39027	Hypothetical protein - <i>Homo sapiens</i> (Human), 456 aa (fragment).	6..461 1..456	456/456 (100%) 456/456 (100%)	0.0
Q9BX97	PV1 protein - <i>Homo sapiens</i> (Human), 442 aa.	20..461 1..442	442/442 (100%) 442/442 (100%)	0.0
Q9BZD5	Fenestrated-endothelial linked structure protein - <i>Homo sapiens</i> (Human), 442 aa.	20..461 1..442	441/442 (99%) 441/442 (99%)	0.0

BAC04681	CDNA FLJ38711 fis, clone KIDNE2003507, highly similar to <i>Homo sapiens</i> PVI protein (PLVAP) mRNA - <i>Homo sapiens</i> (Human), 437 aa.	20..461 1..437	436/442 (98%) 436/442 (98%)	0.0
Q91VC4	MECA32 (Similar to PLASMALEMMA vesicle associated protein) - <i>Mus musculus</i> (Mouse), 438 aa.	20..461 1..438	273/442 (61%) 351/442 (78%)	e-156

PFam analysis predicts that the NOV32a protein contains the domains shown in Table 32E.

Table 32E. Domain Analysis of NOV32a			
Pfam Domain	NOV32a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 33.

The NOV33 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 33A.

Table 33A. NOV33 Sequence Analysis			
	SEQ ID NO: 157	1482 bp	
NOV33a, CG52919-01 DNA Sequence	CCAGGCGCTGGCCGTGGTGTGATTCTGTCTAGGCGCTGGCGCGGCAGCGGCGGTGACGG CTGCGGCCCCGCTCCCTCTACCCGGCCGGACCCGGCTCTGCCCCGCGCCCAAGCCCCAC CAAGCCCCCGCCCTCCCGCCGCGGTCCAGCCAGGGCGCGCCGCAACCAGCACCATTG CGCCCCGTAGCCCTGTCTGCTCCTGCCCCTCGCTGTGGCGCTCCTGGCTCACGGACTCTCT TTAGAGGCCCCAACCGTGGGGAAGGACAAGCCCCAGGCATCGAGGAGACAGATGGCGAG CTGACAGCAGCCCCACACCTGAGCAGCCAGAACGAGGCGTCCACTTTGTCAACAGCC CCCACCTTGAAGCTGCTCAACCACCACCCGCTGCTTGAGGAATTCTACAAGAGGGGCTG GAAAAGGGAGATGAGGAGCTGAGGCCAGCACTGCCCTTCCAGCCTGACCCACCTGCACCC TTCACCCCAAGTCCCCCTTCCCGCCTGGCCAACCAGGACAGCCGCCCTGTCTTTACCAGC CCCACTCCAGCCATGGCTGCGGTACCACTCAGCCCCAGTCCAAGGAGGGACCCCTGGAGT CCGGAGTCAGAGTCCCCCTATGCTTCGAATCACAGCTCCCCCTACCTCCAGGGCCAGCATG GCAGTGCCCAACCTAGGCCAGGGGAGATAGCCAGCACTACACCCCCAGCAGAGCCTGG ACACCAACCAAGAGGGTCTTGAGACATGGGAAGGCGTGGGTGTCAGAGGTTGTGTCC CAGGGCGCAGGGATCGGGATCCAGGGGACCATCACCCTCCTCCACAGCTTCAGGAGATGAT GAGGAGACCACCACTACCACCACCATCATCACCACCACCATCACCACAGTCCAGACACCA GGTCAGCTACCTGCTGGCTTGAGATGTGGAATGGGGATGGGGAGGCTGCGGGGCCCC TAAAAGCCTGTCTCTGACACTGTGCCAGCCTGCCCTGCCCTTTGGCACCAGGGCCAGCC TGCAGGAGGCATGTAGATTGGACCCAGATAGACCTGAGCTCAAATCCTGATTCTTCAGCC AAGTACAGTGGCTCATGCTGTAATCCAGCACTTTGGGAGGCAGAGGCCAGTGGATCAT CTGAGGTGAGGATTCAAGACCCTCTGGCCAACATGGCGAAACACCATCTCTACTAAAA ATACAAAAATGAGCCGGGCATGGTGGTGGGCACCTGTAATCCAGCTACTCGGGAGGCTG AGGCAGGAGAATCACTCAAACCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGCACCAT TGCACTCCAGCCTGGGCAACAGAGCGAGACTCTGTCTCAAAAAAGAAAAATCTTGATTTC TTCCAACATAACATGACCCTAGGAATTCTATTAAACATCTCATCTCTGAGCCTCATCTG		

233

	SEQ ID NO: 161	2127 bp	
NOV33c, CG52919-03 DNA Sequence	CCAGGCGCTGGCCGTGGTGCTGATTCTGTGTCAGGCGCTGGCGGCGGCAGCGCGGTGACGG CTGCGGCCCCGCTCCCTCTACCCGGCCGGACCCGGCTCTGCCCCGCGCCCAAGCCCCAC CAAGCCCCCGCCCTCCCGCCGCGGTCCAGCCCAGGGCGCGGCCGCAACCAGCACCATG CGCCCGGTAGCCCTGCTGCTCCTGCCCTCGCTGCTGGCGCTCCTGGCTCAGGACTCTCT TTAGAGGCCCCAACCGTGGGGAAGGACAAGCCCCAGGCATCGAGGAGACAGATGGCGAG CTGACAGCAGCCCCACACCTGAGCAGCCAGAACGAGGCGTCCACTTTGTCAACAGCC CCCACCTTGAAGCTGCTCAACCACCACCCGCTGCTTGAGGAATTCCTACAAGAGGGGCTG GAAAAGGGAGATGAGGAGCTGAGGCCAGCACTGCCCTTCCAGCCTGACCCACCTGCACCC TTCACCCCAAGTCCCCTTCCCGCCTGGCCAACCAGGACAGCCGCCCTGTCTTTACCAGC CCCCTCCAGCCATGGCTGCGGTACCCACTCAGCCCCAGTCCAGGAGGGAGCCCTGGAGT CCGGAGTCAGAGTCCCCTATGCTTCGAATCACAGCTCCCCTACCTCCAGGGCCAGCATG GCAGTGCCACCCCTAGGCCACAGGGAGATAGCCAGCACTACACCCCGCAGAGAGCCTGG ACACCAACCCAAGAGGTCCTGGAGACATGGGAAGGCCGTGGGTTGCAGAGGTTGTGTCC CAGGGCGCAGGGATCGGGATCCAGGGACCATCACCTCCTCCACAGCTTCAGGAGATGAT GAGGAGACCACCACTACCACCACCATCATCACCACCACCATCACCACAGTCCAGACACCA GGCCCTTGTAGCTGGAATTTCTCAGGCCCAGAGGGCTCTCTGGACTCCCCTACAGACCTC AGCTCCCCCACTGATGTTGGCTGGACTGCTTCTTCTACATCTGTCTACCCCTGGCTAT GGCGTGAAATCAAGGTCCAGAATATCAGCCTCCGGGAAGGGGAGACAGTACTGTGGAA GGCCTGGGGGGGCTGACCCACTGCCCTGGCCAACCAGTCTTCTCTGCTCGGGGCCAA GTCATCCGCAGCCCCACCCACCAAGCGGCCCTGAGGTTCCAGAGCCTCCCGCCACCGGCT GGCCCTGGCACCTTCCATTTCATTACCAAGCCTATCTCCTGAGCTGCCACTTTCGCCGT CGTCCAGCTTATGAGATGTGACTGTCAACAGCCTCCACCCAGGGGTAGTGCCCGGTTTC CATTGTGCCACTGGCTACCAGCTGAAGGGCGCCAGGCATCTCACCTGTCTCAATGCCACC CAGCCCTTCTGGGATTCAAAGGAGCCCGTCTGCATCGCTGCTTGGCGGGAGTGATCCGC AATGGCACCACCGGCCGCATCGTCTCTCCAGGCTTCCCGGGCACTACAGCAACAACCTC ACCTGTCACTGGCTGCTTGAGGCTCCTGAGGGCCAGCGGCTACACCTGCCTTTGAGAAG GTTTCCCTGGCAGAGGATGATGACAGGCTCATCATTGCAATGGGGACAACGTGGAGGCC CCACCAGTGATGATTCTATGAGGTGGAATACCCGCCCGCCCCGCCCTACAACCGC ATTACCATAGAGTCAGCGTTTGACAATCCAACCTACGAGACTGGAGAGACGAGAGAATAT GAAGTCTCCATCTAGGTGGGGGCAGTCTAGGGAAGTCAACTCAGACTTGCAACCACAGTCC AGCAGCAAGGCTCCTTGCTTCTGCTGTCCCTCCACCTCCTGTATATACCACCTAGGAGG AGATGCCACCAAGCCCTCAAGAAGTTGTGCCCTTCCCGCCTGCGATGCCACCATGGCC TATTTCTTGGTGTCAATTGCCCACTTGGGGCCCTTGCATTGGGCCATGTACAGGGGCAT CTACCTGTGGGGAAGAACATAGCTGGGAGCACAAGCTTCAACAGCCAGCATTCCTTGAGC CTCCTTCATGGCCCTGGGACCAGCCTGGGGAACACANTTAGGCAGGAGCAGGGAGTTACC TTGTTTCACATGACCACCAACCATTCC		
	ORF Start: ATG at 178		ORF Stop: TAG at 1753
	SEQ ID NO: 162	525 aa	MW at 56462.7kD
NOV33c, CG52919-03 Protein Sequence	MRPVALLLLPSLLALLAHGLSLEAPTVGKGQAPGIEETDGELTAAPTPEQPERGVHFVTT APTLKLLNHHPLLEEFLEQGLEKGDLELRPALFPQDPDPAPFTPSPLPRLANQDSRPVFT SPTPAMAAVPTQPQSKEGPWSPESPEMLRITAPLPPGSPMAVPTLGPGEIASTTPPSRA WPTPQEGPGDMGRPWWAEVVSQAGIGIQGTITSSASGDDEETTTTTITTTITVTQT PGPCSWNFSGPEGSLDSPDLSSPTDVGLDCFFYISVYPGYGVEIKVQNISLREGETVTV EGLGGPDPLPLANQSFLLRQVIRSPTHQAALRFQSLPPPAGPGTFHFHYQAYLLSCHFP RRPAYGDVTVTSLHPGGSARFHCATGYQLKGARHLTCLNATQPFWDSKEPVCIAACGGVI RNGTTGRIVSPGFPGNYSNNLTCHWLLEAPEGQRLHLHFKEVSLAEDDDRLLIIRNGDNVE APPVYDSYEVEYPPRPRPNRITIEAFDNPTYETGETREYEVSI		
	SEQ ID NO: 163	1988 bp	
NOV33d, CG52919-04 DNA Sequence	CCAGGCGCTGGCCGTGGTGCTGATTCTGTGTCAGGCGCTGGCGGCGGCAGCGCGGTGACGG CTGCGGCCCCGCTCCCTCTACCCGGCCGGACCCGGCTCTGCCCCGCGCCCAAGCCCCAC CAAGCCCCCGCCCTCCCGCCGCGGTCCAGCCCAGGGCGCGGCCGCAACCAGCACCATG CGCCCGGTAGCCCTGCTGCTCCTGCCCTCGCTGCTGGCGCTCCTGGCTCAGGACTCTCT TTAGAGGCCCCAACCGTGGGGAAGGACAAGCCCCAGGCATCGAGGAGACAGATGGCGAG CTGACAGCAGCCCCACACCTGAGCAGCCAGAACGAGGCGTCCACTTTGTCAACAGCC CCCACCTTGAAGCTGCTCAACCACCACCCGCTGCTTGAGGAATTCCTACAAGAGGGGCTG GAAAAGGGAGATGAGGAGCTGAGGCCAGCACTGCCCTTCCAGCCTGACCCACCTGCACCC		

	TTCACCCCAAGTCCCTTCCCGCCTGGCCAACCAGGACAGCCGCCCTGTCTTTACCAGC CCCACTCCAGCCATGGCTGCGGTACCACTCAGCCCCAGTCCAAGGAGGGACCCTGGAGT CCGGAGTCAGAGTCCCCTATGCTTCAATCACAGTCCCCTACCTCCAGGGCCAGCATG GCAGTGGCCACCCTAGGCCAGGGGAGATAGCCAGCACTACACCCCCAGCAGAGCCTGG ACACCAACCCAAGAGGGTCTGGAGACATGGGAAGGCCGTGGGTGTCAGAGGTTGTGTCC CAGGGCGCAGGGATCGGGATCCAGGGGACCATCACCTCCTCCACAGCTTCAGGAGATGAT GAGGAGACCACCACTACCACCACCATCATCACCACCACCATCACCACAGTCCAGACACCA GGCCCTTGTAGCTGGAATTTCTCAGGCCAGAGGGCTCTCTGGACTCCCCTACAGACCTC AGCTCCCCCACTGATGTTGGCTGGACTGCTTCTTCTACATCTCTGTCTACCCTGGCTAT GGCGTGGAAATCAAGGTCCAGAATATCAGCCTCCGGGAAGGGGAGACAGTGAAGTGTGGAA GGCCTGGGGGGCCTGACCACTGCCCTGGCCAACCAGTCTTCTGTCTGCGGGGCCAA GTATCCGCAGCCCCACCCACCAAGCGGCCCTGAGGTTCCAGAGCCTCCCGCCACCGGCT GGCCCTGGCACCTTCCATTTCATTACCAAGCCTATCTCTGAGCTGCCACTTTCCCGT CGTCCAGCTTATGGAGATGTGACTGTCAACAGCCTCCACCCAGGGGGTAGTGGCCGTTT CATTGTGCCACTGGCTACCAGCTGAAGGGCGCCAGGCATCTCACCTGTCTCAATGCCACC CAGCCCTTCTGGGATTCAAAGGAGCCCGTCTGCATCGCTGCTTGGCGCGAGTGATCCGC AATGGCACCACCGCCCGCATCGTCTCTCAGGCTTCCCGGGCAACTACAGCAACAACTC ACCTGTCACTGGCTGCTTGAAGGCTCCTGAGGGCCAGCGGCTACACCTGCCTTTGAGAAG GTTTCCCTGGCAGAGGATGATGACAGGCTCATCATTCGCAATGGGGACAACGTGGAGGCC CCACCACTGTATGATTCTATGAGGTGGAATACCCGCCCCGCCCCGCCCTACAACCGC ATTACCATAGAGTCAGCGTTTGACAATCAACTTACGAGACTGGAGAGACGAGAGAATAT GAAGTCTCCATCTAGGTGGGGCAGTCTAGGGAAGTCAACTCAGACTTGACCACAGTCC AGCAGCAAGGCTCCTTGTCTCTGCTGTCCCTCCACCTCCTGTATATACCACCTAGGAGG AGATGCCACCAAGCCACTTTGTACATGTAATGTATTATATGGGGTCTGGGCTCCAGCCAG AGAACAATCTTTATTTCTGTTGTTTCCTATTAAATGGTGTTTTGGAAAAA AAAAAA		
	ORF Start: ATG at 178		ORF Stop: TAG at 1753
	SEQ ID NO: 164	525 aa	MW at 56462.7kD
NOV33d, CG52919-04 Protein Sequence	MRPVALLLLPSLLALLAHGLSLEAPTVGKGQAPGIEETDGETAAPTPEQPERGVHFTT APTLKLLNHHPLLEEFQEGLEKGDDELRPALPFQPDPPAPFTPSPLPRLANQDSRPVFT SPTPAMAAVPTQPSKEGPWSPESPLRITAPLPGPSMAVPTLGPGEIASTTPPSRA WPTPTQEGPGDMGRPWVAEVVSQAGIGIQTITSSTASGDDEETTTTTTITTTITTVQT PGPCSWNFSGPEGLDSDPTDLSSPTDVGLDCFFYISVYPGYGVEIKVQNISLREGETVT EGLGGPDPLPLANQSFLLRGQVIRSPTHQAALRFQSLPPPAGPGTFHFHYQAYLLSCHFP RRPAYGDVTVTSLHPGGSARFHCATGYQLKGARHLTCLNATQPFWDSKEPVCIAACGGVI RNGTTGRIVSPGFPNYSNNLTCHWLEAPEGQRLHLHFEKVSLEADDDRLLIRNGDNVE APPVYDSYEVEYPPRPRPNRITIESAFDNPTYETGETREYEVSI		
	SEQ ID NO: 165	2143 bp	
NOV33e, CG52919-05 DNA Sequence	CCAGGCGCTGGCCGTGGTGTGATTCTGTCAAGGCGCTGGCGCGCGCAGCGCGGTGACGG CTGCGGCCCCGCTCCCTCTACCCGGCCGGACCCGGCTCTGCCCCGCGCCCAAGCCCCAC CAAGCCCCCGCCCTCCCGCCGCGGTCCAGCCAGGGCGCGCGCAACCAGCACCATG CGCCCGGTAGCCCTGCTGCTCCTGCCCTCGCTGCTGGCGCTCCTGGCTACGGACTCTCT TTAGAGGCCCCAACCGTGGGGAAGGACAAGCCCCAGGCATCGAGGAGACAGATGGCGAG CTGACAGCAGCCCCACCTGAGCAGCCAGAACGAGGCGTCCACTTTGTCAACAACAGCC CCACCTTGAAGCTGCTCAACCACCAACCCGCTGCTTGAGGAATTCCTACAAGAGGGCTG GAAAAGGGAGATGAGGAGCTGAGGCCAGCACTGCCCTTCAGCCTGACCCACCTGCACCC TTCACCCCAAGTCCCCTTCCCGCCTGGCCAACCAGGACAGCCGCCCTGTCTTTACCAGC CCCACTCCAGCCATGGCTGCGGTACCACTCAGCCCCAGTCCAAGGAGGGACCCTGGAGT CCGGAGTCAGAGTCCCCTATGCTTCAATCACAGTCCCCTACCTCCAGGGCCAGCATG GCAGTGGCCACCCTAGGCCAGGGGAGATAGCCAGCACTACACCCCCAGCAGAGCCTGG ACACCAACCCAAGAGGGTCTGGAGACATGGGAAGGCCGTGGGTGTCAGAGGTTGTGTCC CAGGGCGCAGGGATCGGGATCCAGGGGACCATCACCTCCTCCACAGCTTCAGGAGATGAT GAGGAGACCACCACTACCACCACCATCATCACCACCACCATCACCACAGTCCAGACACCA GGCCCTTGTAGCTGGAATTTCTCAGGCCAGAGGGCTCTCTGGACTCCCCTACAGACCTC AGTCCCCCACTGATGTTGCCTGGACTGCTTCTTCTACATCTCTGTCTACCTGGCTAT GGCGTGGAAATCAAGGTCCAGAATATCAGCCTCCGGGAAGGGGAGACAGTGAAGTGTGGAA GGCCTGGGGGGCCTGACCACTGCCCTGGCCAACCAGTCTTCTGTCTGCGGGGCCAA GTATCCGCAGCCCCACCCACCAAGCGGCCCTGAGGTTCCAGAGCCTCCCGCCACCGGCT		

	GGCCCTGGCACCTTCCATTTCATTACCAAGCCTATCTCCTGAGCTGCCACTTTCCCCGT CGTCCAGCTTATGAGACTGTGACTGTACCAGCCTCCACCCAGGGGGTAGTGGCCGCTTC CATTGTGCCACTGGCTACCAGCTGAAGGGCGCCAGGCATCTCACCTGTCTCAATGCCACC CAGCCCTTCTGGGATTCAAAGGAGCCCCGTCTGCATCGCTGCTTCGCGCGGAGTGATCCGC AATGGCACCAACGGCCGCATCGTCTCTCCAGGCTTCCCAGGGCACTACAGCAACAACCTC ACCTGTCACTGGCTGCTTGAGGCTCCTGAGGGCCAGCGGTACACCTGCACCTTTGAGAAG GTTTCCCTGGCAGAGGATGATGACAGGCTCATCATTCGCAATGGGGACAACGTGGAGGCC CCACCAGTGGGAAAAAGCTCCCTGCAGCTGCCCGCCCCCGCCCCGCCCTACAACCGC ATTACCATAGAGTCAGCGTTTGACAATCCAACCTACGAGACTGGATCTCTTTCTTTGCA GGAGACGAGAGAATATGAAGTCTCCATCTAGGTGGGGCAGCTAGGGAAGTCAACTCAG ACTTGACCAACAGTCCAGCAGCAAGGCTCCTTGCTTCTGCTGTCCCTCCACCTCCTGTA TATACCACCTAGGAGGAGATGCCACCAAGCCCTCAAGAAGTTGTGCCCTTCCCCGCTGC GATGCCACCATGGCTATTTTCTTGGTGTATTGCCCACTTGGGGCCCTTGCATTGGGC CATGTACAGGGGGCATCTACCTGTGGGGAAGAACATAGCTGGGAGCACAAGCTTCAACAG CCAGCATTCCTTGAGCCTCCTTCATGGCCCTGGGACCAGCCTGGGGAACACANTTAGGCA GGAGCAGGAGTTACCTTGTTTACATGACCAACCAACCATTCC		
	ORF Start: ATG at 178		ORF Stop: TGA at 1756
	SEQ ID NO: 166	526 aa	MW at 56252.6kD
NOV33e, CG52919-05 Protein Sequence	MRPVALLLLPSLLALLAHGLSLEAPTVGKQAPGIEETDDELTAAPTPEQPERGVHFVTT APTLKLLNHHPLLEEFLEQGLEKGDLELRPALFPQDPPAPFTPSPLPRLANQDSRPVFT SPTPAMAAVPTQPQSKGPPSPSESPMLRITAPLPPGSPMAVPTLGPGEIASTTPPSRA WTPTQEGPGDMGRPWVAEVVSQAGIGIQGTITSSTASGDDEETTTTTTIITTTITVTQ PGPCSWNFGSGPEGLSDSPTDLSSPTDVGLDCFFYISVYPGYGVEIKVQNISLREGETVT EGLGGPDLPLANQSFLLRGQVIRSPHQALRFQSLPPAGPGTFHFHYQAYLLSCHFP RRPAYGDVTVTSLHPGGSARFHCATGYQLKGARHLTCLNATQPFWDSKEPVCIAACGGVI RNGTTGRIVSPGFPNGYSNNLTCHWLLEAPEGQRLHLHFEKVSLEADDDRLIIRNGDNVE APPVGKSSLQLPRPRPRPNRITIESAFDNPTYETGSLSFAGDERI		
	SEQ ID NO: 167	1694 bp	
NOV33f, CG52919-06 DNA Sequence	CAGGGCGCGGCCCAACCAGCACCATCGCGCCGCTAGCCCTGCTGCTCCTGCCCTCGCTG CTGGCGCTCCTGGCTCACGACTCTCTTTAGAGGCCCCAACCGTGGGGAAGGACAAGCC CCAGGCATCGAGGAGACAGATGGCGAGCTGACAGCAGCCCCACACCTGAGCAGCCAGAA CGAGGCGTCCACTTTGTCAACAACAGCCCCACCTTGAAGCTGCTCAACCACCACCCGCTG CTTGAGGAATTCCTACAAGAGGGGCTGGAAGGGAGATGAGGAGCTGAGGCCAGCACTG CCCTTCCAGCCTGACCCACCTGCACCCCTTCACCCCAAGTCCCCTTCCCCGCTGGCCAAC CAGGACAGCCGCCCTGTCTTTACCAGCCCCACTCCAGCCATGGCTGCGGTACCCACTCAG CCCCAGTCCAAGGAGGGACCCTGGAGTCCGGAGTCAGAGTCCCTATGCTTCGAATCACA GCTCCCTACCTCCAGGGCCAGCATGGCAGTGCCACCCCTAGGCCCAGGGGAGATAGCC AGCACTACACCCCCCAGCAGAGCCTGGACACCAACCCAAGAGGGTCTTGAGACATGGGA AGGCCGTGGGTTGCAGAGGTTGTGTCCAGGGCGCAGGGATCGGGATCCAGGGGACCATC ACCTCCTCCACAGCTTCAGGAGATGATGAGGAGACCACCACTACCACCACCATCATCACC ACCACCATCACCACAGTCCAGACACCAGGCCCTTGTAGCTGGAATTTCTCAGGCCAGAG GGCTCTCTGGACTCCCCCTACAGACCTCAGCTCCCCCACTGATGTTGGCTGGACTGCTTC TTCTACATCTCTGTCTACCCTGGCTATGGCGTGGAATCAAGGTCCAGAATATCAGCCTC CGGGAAGGGGAGACAGTACTGTGGAAGGCCTGGGGGGGCTGACCCACTGCCCTGGCC AACCAGTCTTTCTGCTGCGGGGCCAAGTCATCCGAGCCCCACCCACCAAGCGGCCCTG AGGTTCCAGAGCCTCCCGCCACCGCTGGCCCTGGCACCTTCCATTTCATTACCAAGCC TATCTCCTGAGCTGCCACTTCCCCGTCGTCAGCTTATGGAGATGTGACTGTACCAGC CTCCACCCAGGGGGTAGTGCCCGCTTCCATTGTGCCACTGGCTACCAGCTGAAGGGCGCC AGGCATCTCACTGTCTCAATGTCAACCAGCCCTTCTGGGATTCAAAGGAGCCCGTCTGC ATCGCTGCTTGGCGGCGAGTGATCCGCAATGCCACCAACCGGCCGATCTCTCCAGGC TTCCCGGGCAACTACAGCAACAACCTCACCTGTCACTGGCTGCTTGAGGCTCCTGAGGGC CAGCGGCTACACCTGCACCTTGAGAAGGTTTCCCTGGCAGAGGATGATGACAGGCTCATC ATTGCAATGGGGACAACGTGGAGGCCCCACAGTGTATGATTCTATGAGGTGGAATAC CTGCCCATTGAGGGCCTGCTCAGCTCTGGCAACACTTCTTTGTTGAGCCCCGCCCCGC CCCCGCCCTACAACCGCATTACCATAGAGTCAGCGTTTGACAATCCAACCTACGAGACT GGATCTCTTCCCTTGACGAGACGAGAGAATATGAAGTCTCCATCTAGGTGGGGCAGT CTAGGGAAGTCAAC		

	ORF Start: ATG at 25		ORF Stop: TGA at 1654
	SEQ ID NO: 168	543 aa	MW at 58351.0kD
NOV33f, CG52919-06 Protein Sequence	MRPVALLLLPSLLALLAHGLSLEAPTVGKGQAPGIEETDGELTAAPTPEQPERGVHFTT APTLKLLNHHPLLEEFLOEGLEKGEELRPALPFQPDPPAPFTPSPLRLANQDSRPVFT SPTPMAAAVPTQPQSKEGPWSPSESEPMRLITAPLPPGPSMAVPTLGPGEIASTTPPSRA WTPTQEGPGDMGRPWVAEVSQAGIGIQGTITSSTASGDDEETTTTTTITTTITTVQT PGPCSWNFSGPEGLSDSPTDLSSPTDVGLDCFFYISVYPGYGVEIKVQNISLREGETVTV EGLGGPDPLPLANQSFLLRQVIRSPTHQAALRFQSLPPPAGPGTFHFHYQAYLLSCHFP RRPAYGDVTVTSLHPGGSARFHCATGYQLKGARHLTCLNVTQPFWDSKEPVCIAACGGVI RNATTGRIVSPGFPNYSNNLTCHWLLEAPEGQRLHLHFEKVSLAEDDDRLLIIRNGDNVE APPVYDSYEVEYLPFIEGLSSGKHFFVEPRPRPRPNRITIESAFDNPTYETGSLSLAGD ERI		
	SEQ ID NO: 169	1482 bp	
NOV33g, CG52919-01 DNA Sequence	CCAGGCGCTGGCCGTGGTGTGATTCTGTCTAGGCGCTGGCGCGGCAGCGCGGTGACGG CTGCGGCCCCGCTCCCTCTACCCGGCCGACCCGCTCTGCCCGCGGCCAAGCCCCAC CAAGCCCCCGCCCTCCCGCCGGTCCCAGCCAGGGCGCGGCCGCAACCAGCACCATG CGCCCGGTAGCCCTGCTGCTCCTGCCCTCGCTGCTGGCGCTCTGGCTCACGGACTCTCT TTAGAGGCCCAACCGTGGGGAAGGACAAGCCCCAGGCATCGAGGAGACAGATGGCGAG CTGACAGCAGCCCCACACCTGAGCAGCCAGAACGAGGCGTCCACTTTGTACACAGCC CCCACCTTGAAGCTGCTCAACCACCACCCGCTGCTTGAGGAATTCTTACAAGAGGGGCTG GAAAAGGGAGATGAGGAGCTGAGGCCAGCACTGCCCTTCCAGCCTGACCCACCTGCACCC TTCACCCCAAGTCCCTTCCCGCCTGGCCAACCAGGACAGCGCCCTGTCTTTACCAGC CCCACTCCAGCCATGGCTGCGGTACCCACTCAGCCCCAGTCCAAGAGGGACCTGGAGT CCGGAGTCAGAGTCCCTATGCTTGAATCACAGTCCCTACCTCCAGGGCCAGCATG GCAGTCCCCACCCTAGGCCAGGGGAGATAGCCAGCACTACACCCCCAGCAGAGCCTGG ACACCAACCCAGAGGGTCTTGAGACATGGGAAGGCCGTGGGTTGCAGAGGTTGTGTCC CAGGGCGCAGGATCGGGATCCAGGGACCACCTCCTCCACAGCTTCAGGAGATGAT GAGGAGACCACCACTACCACCACCATCATCACCACCACCATCACCACAGTCCAGACACCA GGTCAGCTACCTGCTGGCTTGAGATGTGGAATGGGGATGGGGAGGCTGCGGGGCCCC TAAAAGCCTGTCTCTGACACTGTGCCAGCCTGCCCTGCCCTTTGGCACCAAGGGCCAGCC TGCAGGAGGCATGTAGATTGGACCCAGATAGACCTGAGCTCAAACTCTGATTCTTCAGCC AAGTACAGTGGCTCATGCCTGTAATCCAGCACTTTGGGAGGCAGAGCCAGTGGATCAT CTGAGGTACAGGAGTTCAAGACCCCTCCTGGCCAACATGGCGAAACACCTCTCTACTAAAA ATACAAAAATGAGCCGGGCATGGTGGTGGGCACCTGTAATCCAGCTACTCGGGAGGCTG AGGCAGGAGAATCACTCAACCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATTGCACCAT TGCACTCCAGCCTGGGCAACAGAGCGAGACTCTGTCTCAAAAAAGAAAAATCTTGATTC TTCCAATATAACATGACCCTAGGAATTCTATTTAACATCTCATCTCTGAGCCTCATCTG TAAAATGGCAATAAGAAAAATAAACTCTGGCTAGAAAAAAA		
	ORF Start: ATG at 178		ORF Stop: TAA at 961
	SEQ ID NO: 170	261 aa	MW at 27471.8kD
NOV33g, CG52919-01 Protein Sequence	MRPVALLLLPSLLALLAHGLSLEAPTVGKGQAPGIEETDGELTAAPTPEQPERGVHFTT APTLKLLNHHPLLEEFLOEGLEKGEELRPALPFQPDPPAPFTPSPLRLANQDSRPVFT SPTPMAAAVPTQPQSKEGPWSPSESEPMRLITAPLPPGPSMAVPTLGPGEIASTTPPSRA WTPTQEGPGDMGRPWVAEVSQAGIGIQGTITSSTASGDDEETTTTTTITTTITTVQT PGQLPAGLQMWKGGWRLRGP		
	SEQ ID NO: 171	840 bp	
NOV33h, CG52919-07 DNA Sequence	TCCAGGGCGCGCCGCAACCAGCACCATGCGCCCGGTAGCCCTGCTGCTCCTGCCCTGCG TGCTGGCGCTCCTGGCTCACGACTCTCTTTAGAGGCCCAACCGTGGGGAAGGACAAG CCCCAGGCATCGAGGAGACAGATGGCGAGCTGACAGCAGCCCCACACCTGAGCAGCCAG AACGAGGCGTCCACTTTGTACAAACAGCCCCACCTTGAAGCTGCTCAACCACCACCGC TGCTTGAGGAATTCTTACAAGAGGGCCGGAAGGGAGATGAGGAGCTGAGGCCAGCAC TGCCCTTCCAGCCTGACCCACCTGCACCCCTTACCCCAAGTCCCCTTCCCCGCTGGCCA ACCAGGACAGCGCCCTGTCTTTACCAGCCCCACTCCAGCCATGGCTGCGGTACCCACTC AGCCCCAGTCCAAGGAGGACCTGGAGTCCGGAGTCAGAGTCCCCTATGCTTCGAATCA CAGCTCCCCTACCTCCAGGGCCAGCATGGCAGTGCCACCCCTAGGCCAGGGAGATAG CCAGCACTACACCCCCAGCAGAGCCTGGACACCAACCAAGAGGGTCTGGAGACATGG		

	GAAGCCCGTGGGTTGCAGAGGTTGTGTCCAGGGCGCAGGGATCGGGATCCAGGGGACCA TCACCTCTCCACAGCTTCAGGAGATGATGAGGAGACCACCCTACCACCACCATCATCA CCACCACCATCACCACAGTCCAGACACCAGGTGAGTACCTGCTGGCTTGAGATGTGGA AATGGGGATGGGGGAGGCTGCGGGGCCCCATAAAGCCTGTCTCTGACACTGTGCCAGCCA		
	ORF Start: ATG at 27		ORF Stop: TAA at 810
	SEQ ID NO: 172	261 aa	MW at 27455.8kD
NOV33h, CG52919-07 Protein Sequence	MRPVALLLLPSLLALLAHGLSLEAPTGVKGQAPGIEETDDELTAAPTPEQPERGVHFVTT APTLKLLNHHPLLEEFQEGPEKGDEELRPALPFQDPDPAPFTPSPLPRLANQDSRPVFT SPTPMAAAVPTQPQSKGWPSPESPSMLRITAPLPPGSPMAVPTLGPGEIASTTPPSRA WTPTQEGPGDMGRPWWAEVVSQAGIGIGTITSSTASGDDEETTTTTTIITTTITTVQT PGQLPAGLQMWKGGWRLRGP		
	SEQ ID NO: 173	1654 bp	
NOV33i, CG52919-08 DNA Sequence	CACCAGATCTCCACCATGCGCCGGTAGCCCTGTGCTCCTGCCCTCGTGTGGCGCT CCTGGCTCACGGACTCTCTTAGAGGCCCAACCGTGGGGAAAGGACAAGCCCCAGGCAT CGAGGAGACAGATGGCGAGCTGACAGCAGCCCCACACCTGAGCAGCCAGAACGAGGCGT CCACTTTGTCAACAGCCCCACCTTGAAGCTGCTCAACCACCACCCGCTGCTTGAGGA ATTCTACAAGAGGGGCTGGAAAAGGGAGATGAGGAGCTGAGGCCAGCACTGCCCTTCCA GCCTGACCCACCTGCACCCTTCACCCCAAGTCCCTTCCCGCTGGCCCAACGAGCAG CCGCCTGTCTTTACCAGCCCCACTCCAGCCATGGCTGCGGTACCCACTCAGCCCCAGTC CAAGGAGGGACCTGGAGTCCGGAGTCAGAGTCCCCTATGCTTCAATCACAGCTCCCCT ACCTCCAGGGCCAGCATGGCAGTGCCACCCTAGGCCAGGGGAGATAGCCAGCACTAC ACCCCCCAGCAGAGCCTGGACACCAACCAAGAGGGTCTGGAGACATGGGAAGGCCGTG GGTTGCAGAGTGTGTCTCCAGGGCGCAGGGATCGGGATCCAGGGGACCATCACCTCTC CACAGCTTCAGGAGATGATGAGGAGACCACCACTACCACCACCATCATCACCACCACAT CACCACAGTCCAGACACCAGGCCCTTGTAGCTGGAATTTCTCAGGCCAGAGGGTTCTCT GGACTCCCCTACAGACCTCAGCTCCCCCACTGATGTTGGCCTGGACTGCTTCTTCTACAT CTCTGTCTACCTGGCTATGGCGTGGAAATCAAGGTCCAGAATATCAGCCTCCGGGAAGG GGAGACAGTGACTGTGGAAGGCCTGGGGGGGCTGACCACTGCCCTGGCCCAACGAGTC TTTCTGCTGCGGGGCCAAGTCATCCGACAGCCCCACCAACCAAGCGGCCCTGAGGTTCCA GAGCCTCCCGCCACCGGCTGGCCCTGGCACCTTCCATTTCCATTACCAAGCCTATCTCCT GAGCTGCCACTTTCCCGTCTGTCAGCTTATGGAGATGTGACTGTCAACAGCCTCCACCC AGGGGGTAGTCCCGCTTCCATTGTGCCACTGGCTACCAGTGAAGGGCGCCAGGCATCT CACCTGTCTCAATGTCAACAGCCCTTCTGGGATTCAAAGGAGCCGCTGTCATCGCTGC TTGCGGCGGAGTGATCCGCAATGCCACCACCGGCCGCATCGTCTCTCAGGCTTCCCGGG CAACTACAGCAACAACCTCACCTGTCACTGGCTGCTTGAGGCTCCTGAGGGCCAGCGGCT ACACCTGCACTTTGAGAAGGTTTCCCTGGCAGAGGATGATGACAGGCTCATACCTGCGAA TGGGGACAACGTGGAGGCCCCACCACTGTATGATTCTATGAGGTGGAATACCTGCCCAT TGAGGGCCTGCTCAGCTCTGGCAAACTTCTTTGTTGAGCCCCGCCCGCCCCGCCCC CTACAACCGCATTACCATAGAGTCAGCGTTTGACAATCCAACCTTACGAGACTGGATCTCT TTCCCTTGACAGGAGACGAGAGAATACTCGAGGGC		
	ORF Start: ATG at 17		ORF Stop: at 1646
	SEQ ID NO: 174	543 aa	MW at 58351.0kD
NOV33i, CG52919-08 Protein Sequence	MRPVALLLLPSLLALLAHGLSLEAPTGVKGQAPGIEETDDELTAAPTPEQPERGVHFVTT APTLKLLNHHPLLEEFQEGLEKGDDEELRPALPFQDPDPAPFTPSPLPRLANQDSRPVFT SPTPMAAAVPTQPQSKGWPSPESPSMLRITAPLPPGSPMAVPTLGPGEIASTTPPSRA WTPTQEGPGDMGRPWWAEVVSQAGIGIGTITSSTASGDDEETTTTTTIITTTITTVQT PGPCSWNFSGPEGLSDSPDLSSPTDVLDCFFYISVYPGYGVEIKVQNISLREGETVTV EGLGGPDPLPLANQSFLLRGQVIRSPTHQAALRFQSLPPPAGPGTFHFHYQAYLLSCHFP RRPAYGDVTVSLHPGGSARFHCATGYQLKGARHLTCLNVTQPFWDSEKPVIAACGGVI RNATTGRIVSPGFPNGYSNNLTCHWLLEAPEGQRLHLHFEKVSLAEDDDRLLIRNGDNVE APPVYDSYEVYLPVIEGLSSGKHFFVEPRPRPRPNRITIESAFDNPTYETGSLSLAGD ERI		
	SEQ ID NO: 175	1591 bp	
NOV33j, CG52919-09 DNA Sequence	CACCAGATCTCTCTCTTAGAGGCCCAACCGTGGGGAAAGGACAAGCCCCAGGCATCGA GGAGACAGATGGCGAGCTGACAGCAGCCCCACACCTGAGCAGCCAGAACGAGGCGTCCA CTTTGTCAACAGCCCCACCTTGAAGCTGCTCAACCACCACCCGCTGCTTGAGGAATT		

	CCTACAAGAGGGGCTGGAAAAGGGAGATGAGGAGCTGAGGCCAGCACTGCCCTTCCAGCC TGACCCACCTGCACCCCTTACCCCCAAGTCCCCTTCCCCGCTGGCCAACCAGGACAGCCG CCCTGTCTTTACCAGCCCCACTCCAGCCATGGCTGCGGTACCCACTCAGCCCCAGTCCAA GGAGGGACCTGGAGTCCGGAGTCAGAGTCCCCTATGCTTCGAATCAGAGTCCCCTACC TCCAGGGCCCAGCATGGCAGTGCCACCCCTAGGCCAGGGGAGATAGCCAGCACTACACC CCCCAGCAGAGCCTGGACACCAACCAAGAGGGTCTGGAGACATGGGAAGGCCGTGGGT TGCAGAGGTTGTGTCCAGGGCGCAGGGATCGGGATCCAGGGGACCATCACCTCTCCAC AGCTTCAGGAGATGATGAGGAGACCACCACTACCACCACCATCATCACCACCACCATCAC CACAGTCCAGACACCAGGCCCTTGTAGCTGGAATTTCTCAGGCCAGAGGGTTCTCTGGA CTCCCCTACAGACCTCAGTCCCCCACTGATGTGGCCTGGACTGCTTCTTCTACATCTC TGTCTACCCTGGCTATGGCGTGGAAATCAAGGTCCAGAATATCAGCCTCCGGGAAGGGGA GACAGTGACTGTGGAAGGCCTGGGGGGCCTGACCCACTGCCCTGGCCAACCACTCTTT CCTGTGCGGGGCAAGTCATCCGACGCCCCACCCACCAAGCGGCCCTGAGGTTCCAGAG CCTCCCGCCACCGGCTGGCCCTGGCACCTTCCATTTCCATTACCAAGCCTATCTCTGAG CTGCCACTTTCCCGCTCGTCCAGCTTATGGAGATGTGACTGTACCAGCCTCCACCCAGG GGGTAGTGCCCGCTTCCATTGTGCACTGGCTACCAGCTGAAGGGCGCCAGGCATCTCAC CTGTCTCAATGTCAACCCAGCCCTTCTGGGATTCAAAGGAGCCCGTCTGCATCGTGCTTG CGGCGGAGTGATCCGCAATGCCACCACCGGCCGCATCGTCTCTCCAGGTTCCCGGGCAA CTACAGCAACAACCTCACCTGTCACTGGCTGCTTGAGGCTCCTGAGGGCCAGCGGCTACA CCTGCACTTTGAGAAGGTTTCCCTGGCAGAGGATGATGACAGGCTCATCATTGCAATGG GGACAACGTGGAGGCCCCACCACTGTATGATTCTATGAGGTGGAATACCTGCCATTGA GGGCCTGCTCAGCTCTGGCAAACACTTCTTTGTTGAGCCCCGCCCCGCCCCGCCCTA CAACCGATTACCATAGAGTCAGCGTTTGACAATCCAACCTACGAGACTGGATCTCTTTC CCTGTCAGGAGACGAGAGAATACTCGAGGGC		
	ORF Start: at 2		ORF Stop: at 1583
	SEQ ID NO: 176	527 aa	MW at 56714.8kD
NOV33j, CG52919-09 Protein Sequence	TRSLSLEAPTVGKGQAPGIEETDSELTAAPTPEQPERGVHVVTTAPTLLNHHPLLEEF LQEGLEKGDDELRPALPFQPDPPAPFTPSPLRLANQDSRPVFTSPTPAMAAVPTQPQSK EGPWSPESPEPLMRITAPLPPGPSMAVPTLGPGEIASTTPPSRAWPTQEGPGDMGRPWV AEVVSQAGIGIQGTITSSSTASGDDEETTTTIIITTTITVQTPGPCSWNFGSGPEGLD SPTDLSSPTDVGLDCCFFYISVYPGYGVEIKVQNISLREGETVTVEGLGPDPLPLANQSF LLRGQVIRSPHQALRFQSLPPPAGPGTFHFHYQAYLLSCHFRRPAYGDVTVTSLHPG GSARFHCATGYQLKGARHLTCLNVTQPFWDSKEPVCIAACGGVIRNATTGRIVSPGFPGN YSNNLTCHWLLEAPEGQRLHLHFVKVSLAEDDDRLIIRNGDNVEAPPVYDSYEVEYLP GLSSGKHFFVEPRPRPRPNRITIESAFDNPTYETGSLSLAGDERI		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 33B.

Table 33B. Comparison of NOV33a against NOV33b through NOV33j.		
Protein Sequence	NOV33a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV33b	1..242 1..242	166/242 (68%) 166/242 (68%)
NOV33c	1..242 1..242	166/242 (68%) 166/242 (68%)
NOV33d	1..242 1..242	166/242 (68%) 166/242 (68%)
NOV33e	1..242 1..242	166/242 (68%) 166/242 (68%)
NOV33f	1..242 1..242	166/242 (68%) 166/242 (68%)

NOV33g	1..261 1..261	185/261 (70%) 185/261 (70%)
NOV33h	1..261 1..261	184/261 (70%) 184/261 (70%)
NOV33i	1..242 1..242	166/242 (68%) 166/242 (68%)
NOV33j	20..242 4..226	167/223 (74%) 167/223 (74%)

Further analysis of the NOV33a protein yielded the following properties shown in Table 33C.

Table 33C. Protein Sequence Properties NOV33a	
PSort analysis:	0.8200 probability located in outside; 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in lysosome (lumen)
SignalP analysis:	Cleavage site between residues 20 and 21

- A search of the NOV33a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 33D.

Table 33D. Geneseq Results for NOV33a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV33a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB70543	Human PRO13 protein sequence SEQ ID NO:26 - <i>Homo sapiens</i> , 261 aa. [WO200110902-A2, 15-FEB-2001]	1..261 1..261	261/261 (100%) 261/261 (100%)	e-154
AAE15853	Human SEZ6 protein - <i>Homo sapiens</i> , 853 aa. [WO200183552-A2, 08-NOV-2001]	1..242 1..242	242/242 (100%) 242/242 (100%)	e-140
AAU81976	Human secreted protein SECP2 - <i>Homo sapiens</i> , 994 aa. [WO200198353-A2, 27-DEC-2001]	1..242 1..242	242/242 (100%) 242/242 (100%)	e-140

AAB70542	Human PRO12 protein sequence SEQ ID NO:24 - <i>Homo sapiens</i> , 526 aa. [WO200110902-A2, 15- FEB-2001]	1..242 1..242	242/242 (100%) 242/242 (100%)	e-140
AAB70541	Human PRO11 protein sequence SEQ ID NO:22 - <i>Homo sapiens</i> , 525 aa. [WO200110902-A2, 15- FEB-2001]	1..242 1..242	242/242 (100%) 242/242 (100%)	e-140

In a BLAST search of public sequence databases, the NOV33a protein was found to have homology to the proteins shown in the BLASTP data in Table 33E.

Table 33E. Public BLASTP Results for NOV33a				
Protein Accession Number	Protein/Organism/Length	NOV33a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAC33421	Sequence 25 from Patent WO0110902 - <i>Homo sapiens</i> (Human), 261 aa.	1..261 1..261	261/261 (100%) 261/261 (100%)	e-154
CAC33420	Sequence 23 from Patent WO0110902 - <i>Homo sapiens</i> (Human), 526 aa.	1..242 1..242	242/242 (100%) 242/242 (100%)	e-140
CAC33418	Sequence 19 from Patent WO0110902 - <i>Homo sapiens</i> (Human), 525 aa.	1..242 1..242	242/242 (100%) 242/242 (100%)	e-140
CAC33417	Sequence 17 from Patent WO0110902 - <i>Homo sapiens</i> (Human), 525 aa.	1..242 1..242	242/242 (100%) 242/242 (100%)	e-140
CAC33416	Sequence 15 from Patent WO0110902 - <i>Homo sapiens</i> (Human), 994 aa.	1..242 1..242	242/242 (100%) 242/242 (100%)	e-140

PFam analysis predicts that the NOV33a protein contains the domains shown in
5 Table 33F.

Table 33F. Domain Analysis of NOV33a			
Pfam Domain	NOV33a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 34.

The NOV34 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 34A.

Table 34A. NOV34 Sequence Analysis			
	SEQ ID NO: 177	368 bp	
NOV34a, CG55698-01 DNA Sequence	CTGTCCCACTCACCATGGAGAAGATCCTGATCCTCTGCTTGTGCGCCTCTCTGTGGCCT ATGCAGCTCCTGGCCCCGGGGGATCATTATCAACCTGGAGAACGGTGAGCTCTGCATGA ATAGTGCCCACTGTAAGAGCAATTGCTGCCAGCATTCAAGTGCCTGGGCTGGCCCGCT GCACATCCATGGCCAGCGAGAACAGCGAGTGCTCTGTCAAGACGCTCTATGGGATTACT ACAAGTGTCCCTGTGAGCGTGGCCTGACCTGTGAGGGAGACAAGACCATCGTGGGCTCCA TCACCAACACCAACTTTGGCATCTGCCATGACGCTGGACGCTCAAGCAGTGAGACTGCC CACCCT		
	ORF Start: ATG at 15		ORF Stop: TGA at 351
	SEQ ID NO: 178	112 aa	MW at 11953.7kD
NOV34a, CG55698-01 Protein Sequence	MEKILILLVVALSVAYAAPGPRGIIINLENGELCMNSAQCKSNCCQHSSALGLARCTSM SENSECSVKTLGYIYKPCERGLTCEGDKTIVGSITNTNFGICHDAGRSKQ		
	SEQ ID NO: 179	394 bp	
NOV34b, CG55698-02 DNA Sequence	AGCTGTCCCACTCGCCATGGAGAAGATCCTGATCCTCTGCTTGTGCGCCTCTCTGTGGC CTATGCAGCTCCTGGCCCCGGGGGATCATTATCAACCTGACGCTCTATGGGATTACTA CAAGTGTCCCTGTGAGCGTGGCCTGACCTGTGAGGGAGACAAGACCATCGTGGGCTCCAT CACCACACCAACTTTGGCATCTGCCATGACGCTGGACGCTCCAAGCAGTGAGACTGCCC ACCCACTCCACACCTAGCCAGAAATGCTGTAGGCCACTAGGCGCAGGGGCATCTCTCCC CTGCTCCAGCGCATCTCCCGGGCTGGCCACCTCCTTGACCAGCATATCTGTTTTCTGATT GCGCTCTTCAATTAAGGCCTCCTGCAAACCT		
	ORF Start: ATG at 17		ORF Stop: TGA at 230
	SEQ ID NO: 180	71 aa	MW at 7658.9kD
NOV34b, CG55698-02 Protein Sequence	MEKILILLVVALSVAYAAPGPRGIIINLTLYGIYKPCERGLTCEGDKTIVGSITNTNF GICHDAGRSKQ		

- Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 34B.

Table 34B. Comparison of NOV34a against NOV34b.		
Protein Sequence	NOV34a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV34b	1..112	56/112 (50%)
	1..71	56/112 (50%)

Four polymorphic variants of NOV34b have been identified and are shown in Table 41M.

Further analysis of the NOV34a protein yielded the following properties shown in Table 34C.

Table 34C. Protein Sequence Properties NOV34a	
PSort analysis:	0.8200 probability located in outside; 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in lysosome (lumen)
SignalP analysis:	Cleavage site between residues 18 and 19

A search of the NOV34a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 34D.

Table 34D. Geneseq Results for NOV34a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV34a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB54163	Human pancreatic cancer antigen protein sequence SEQ ID NO:615 - <i>Homo sapiens</i> , 131 aa. [WO200055320-A1, 21-SEP-2000]	1..112 20..131	112/112 (100%) 112/112 (100%)	3e-62
AAY91513	Human secreted protein sequence encoded by gene 63 SEQ ID NO:186 - <i>Homo sapiens</i> , 122 aa. [WO200006698-A1, 10-FEB-2000]	28..111 33..114	28/84 (33%) 38/84 (44%)	1e-07
AAY35930	Extended human secreted protein sequence, SEQ ID NO. 179 - <i>Homo sapiens</i> , 121 aa. [WO9931236-A2, 24-JUN-1999]	28..111 33..114	28/84 (33%) 38/84 (44%)	1e-07
AAB62640	Human colipase-like protein-1 (Zclps1) - <i>Homo sapiens</i> , 118 aa. [WO200136466-A2, 25-MAY-2001]	32..111 34..111	26/80 (32%) 36/80 (44%)	3e-07
AAB62648	Human colipase-like protein-1 (Zclps1) fragment - <i>Homo sapiens</i> , 97 aa. [WO200136466-A2, 25-MAY-2001]	32..109 22..97	25/78 (32%) 35/78 (44%)	1e-06

In a BLAST search of public sequence databases, the NOV34a protein was found to have homology to the proteins shown in the BLASTP data in Table 34E.

Table 34E. Public BLASTP Results for NOV34a				
Protein Accession Number	Protein/Organism/Length	NOV34a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P04118	Colipase precursor - <i>Homo sapiens</i> (Human), 112 aa.	1..112 1..112	112/112 (100%) 112/112 (100%)	9e-62
P19090	Colipase precursor - <i>Canis familiaris</i> (Dog), 112 aa.	1..112 1..112	88/112 (78%) 99/112 (87%)	6e-50
P42890	Colipase precursor - <i>Oryctolagus cuniculus</i> (Rabbit), 107 aa.	1..106 1..106	88/106 (83%) 97/106 (91%)	1e-49
Q91XL7	Pancreatic colipase - <i>Spermophilus tridecemlineatus</i> (Thirteen-lined ground squirrel), 111 aa.	3..112 2..111	87/110 (79%) 100/110 (90%)	2e-49
Q9NIT6	Colipase - <i>Sus scrofa</i> (Pig), 112 aa.	1..110 1..110	86/110 (78%) 95/110 (86%)	1e-48

PFam analysis predicts that the NOV34a protein contains the domains shown in

5 Table 34F.

Table 34F. Domain Analysis of NOV34a			
Pfam Domain	NOV34a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Colipase	21..60	32/40 (80%) 40/40 (100%)	5.5e-24
Colipase_C	62..106	41/47 (87%) 45/47 (96%)	3.2e-34

Pancreatic lipase catalyzes the hydrolysis triacylglycerol to fatty acids. These triacylglycerides are present predominantly as an emulsified micelle stabilized by bile acids. Since lipase hydrolyzes the ester linkage of triacylglyceride, the active site must be

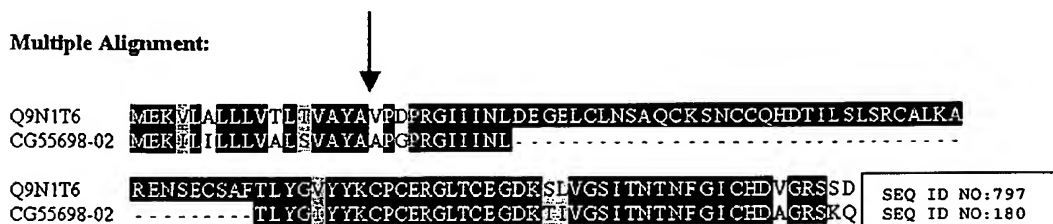
10 positioned at the bile salt-coated water-lipid interface of this micelle. Since the bile salts can

inhibit lipase, colipase is secreted to anchor the lipase to the water-lipid interface so that hydrolysis can occur.

Table 34G shows an alignment of the porcine pancreatic colipase (Q9N1T6; SEQ ID NO:797) with the splice variant NOV34b (CG55698-02; SEQ ID NO:180). The arrow indicates the signal sequence cleavage site. Since the homology between the porcine and human lipases is high, the x-ray crystal structure of the porcine lipase is a suitable comparison for the effects of NOV34b (CG55698-02).

Table 34G. Multiple Alignment of Q9N1T6 and NOV34b (CG55698-02)

Multiple Alignment:



Q9N1T6	MEKVLALLVTLTVAYAVDP	PRGIIINLDEGELCLNSAQCKSNCCO	HTILSLSRCALKA	
CG55698-02	MEKVLALLVALLSVAYAAP	CGPRGIIINL	-----	
Q9N1T6	RENSECSAFTLYGMY	YKPCERGLTCEGDKSLV	GSITNTNFGICHDVGRSSD	SEQ ID NO:797
CG55698-02	-----TLYGMY	YKPCERGLTCEGDKSLV	GSITNTNFGICHDAGRSKQ	SEQ ID NO:180

Figure 2 shows the x-ray crystal structure (1ETH) at a 2.84 Å resolution of porcine lipase (right) with colipase (left)(Hermoso, *et. al*, J. Biol. Chem., 2001, 271:1807-18016).

The tetra ethylene glycol mono-octyl ether inhibitor is shown in the active site of lipase. The deleted sequence found in NOV34b is indicated with hatch marks.

The amino-terminal domain of lipase contains the active site whereas the carboxy-terminal domain binds to colipase. Likewise, colipase possesses a lipase binding domain and a micelle interfacial binding site. The catalytic site of lipase is inaccessible in solution since there is an N-terminal flap which covers the active site, preventing substrate from entering. The colipase additionally serves to stabilize the active form of lipase by binding to the N-terminal flap and thus keeping it in an open, active conformation which allows substrate to enter the lipase active site.

The interfacial binding site of colipase is composed of four hydrophobic fingers (finger1:14-24, finger2:27-39, finger3:47-64, and finger4: 68-90 numbered according to the colipase sequence in Figure 3). In NOV34b, Fingers1, 2 and a portion of 3 are missing suggesting that the splice variant would be less adept at binding the micelle interface.

- Of the 8 polar interactions (includes hydrogen bonds and salt bridges) between lipase and colipase, 5 of bind to the C-terminal region of lipase and the remainder bind to the N-terminal flap. Of these, only one of the 5 bonds NOV34b:C-terminal bonds is missing, but all three of the NOV34b:N-terminal flap bonds missing. Of the 17
- 5 colipase:lipase van der Waals contacts, 4 of these contact the N-terminal flap and the remainder bond to the C-terminal domain. For NOV34b, 11 of the 13 van der Waals contacts to the lipase C-terminal domain and none of the N-terminal flap contacts are present. Of the 4 bridging water contacts at the colipase:lipase C-terminal binding site, 2 are lost in NOV34b.
- 10 The splice variant NOV34b retains most of the binding sites to the C-terminal of lipase, but are missing half of the micelle interfacial binding domain and the entire N-terminal flap binding site. NOV34b may still bind to lipase, but may not anchor it to the micelle interface very well and would not be able to stabilize the open, active formation of lipase (since it cannot bind the N-terminal flap). Thus, it is possible that NOV34b may
- 15 compete for binding with the normal, lipase-activating form of colipase to lipase. Since the NOV34b lipase complex fails to position the N-terminal flap away from the active site of lipase and thus prevents substrate binding, NOV34b may be considered to be a competitive inhibitor of the lipase enzymatic activity.

Example 35.

- 20 The NOV35 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 35A.

Table 35A. NOV35 Sequence Analysis			
	SEQ ID NO: 181	7286 bp	
NOV35a, CG55832-01 DNA Sequence	GAATTCGCTAGAGCCCTAGAGCCCCAGCAGCACCCAGCCAAACCCACCTCCACCATGGGG GCCATGACTCAGCTGTTGGCAGGTGTCTTCTTGCTTTCCTTGCCCTCGCTACCGAAGGT GGGGTCCTCAAGAAAGTCATCCGGCACAAGCGACAGAGTGGGGTGAACGCCACCCTGCCA GAAGAGAACCAGCCAGTGGTGTTTAACCACGTTTACAACATCAAGCTGCCAGTGGGATCC CAGTGTTTCGGTGGATCTGGAGTCAGCCAGTGGGGAGAAAGACCTGGCACCGCCTTCAGAG CCCAGCGAAAGCTTTCAGGAGCACACAGTAGATGGGGAAAACCAGATTGTCTTCACACAT CGCATCAACATCCCCCGCGGCCTGTGGCTGTGCCGAGCCCTGATGTTAAGGAGCTG CTGAGCAGACTGGAGGAGCTGGAGAACCCTGGTGTCTTCCCTGAGGGAGCAATGTACTGCA GGAGCAGGCTGCTGTCTCCAGCCTGCCACAGGCCGCTTGGACACCAGGCCCTTCTGTAGC GGTTCGGGGCAACTTCAGCACTGAAGGATGTGGCTGTGTCTGCGAACCTGGCTGGAAAGGC CCCAACTGCTCTGAGCCCCGAATGTCCAGGCAACTGTCACCTTCGAGGCCGCTGCATTGAT GGGCAGTGCATCTGTGACGACGGCTTCAGGGCGAGGACTGCAGCCAGCTGGCTTGCCCC AGCGACTGCAATGACCAGGGCAAGTGCGTGAATGGAGTCTGCATCTGTTTCGAAGGCTAC GCGGCTGACTGCAGCCGTGAAATCTGCCAGTGCCTGCAGTGAGGAGCACGGCACATGT GTAGATGGCTTGTGTGTGTGCCACGATGGCTTTCAGGCGATGACTGCAACAAGCCTCTG TGTCTCAACAATTGCTACAACCGTGGACGATGCGTGGAGAATGAGTGCCTGTGTGATGAG GGTTCACGGGCGAAGACTGCAGTGAGTCTATCTGCCCAATGACTGCTTCGACCGGGGC		

CGCTGCATCAATGGCACCTGCTACTGCGAAGAAGGCTTCACAGGTGAAGACTGCGGGAAA
CCCACCTGCCCACATGCCTGCCACACCCAGGGCCGGTGTGAGGAGGGGCAGTGTGTATGT
GATGAGGGCTTTGCCGGTGTGGACTGCAGCGAGAAGAGGTGTCTGTGCTACTGTCAACAT
CGTGGCCGCTGTGTAGACGGGCGGTGTGAGTGTGATGATGGTTTCACTGGAGCTGACTGT
GGGGAGCTCAAGTGTCCCAATGGCTGCAGTGGCCATGGCCGCTGTGTCAATGGGCAGTGT
GTGTGTGATGAGGGCTATACTGGGGAGGACTGCAGCCAGCTACGGTGCCCAATGACTGT
CACAGTCGGGGCCGCTGTGTGCGAGGGCAAATGTGTATGTGAGCAAGGCTTCAAGGGCTAT
GACTGCAGTGACATGAGCTGCCCTAATGACTGTCACCAGCACGGCCGCTGTGTGAATGGC
ATGTGTGTTTGTGATGACGGCTACACAGGGGAAGACTGCCGGGATCGCCAATGCCCCAGG
GACTGCAGCAACAGGGGCCCTCTGTGTGGACGGACAGTGCCTCTGTGAGGACGGCTTCACC
GGCCCTGACTGTGCAGAACTCTCTGTCCAAATGACTGCCATGGCCAGGGTCGCTGTGTG
AATGGGCAGTGCCTGTGCCATGAAGGATTTATGGGCAAAGACTGCAAGGAGCAAAGATGT
CCCAGTGACTGTGATGGCCAGGGCCGCTGCGTGGACGGCCAGTGCATCTGCCACGAGGGC
TTCACAGGCCTGGACTGTGGCCAGCACTCCTGCCCCAGTGACTGCAACAACTTAGGACAA
TGCGTCTCGGGCCGCTGCATCTGCAACGAGGGCTACAGCGGAGAAGACTGTCTCAGAGGTG
TCTCTCCCAAAGACCTCGTTGTGACAGAAGTGACGGAAGAGACGGTCAACCTGGCTGG
GACAATGAGATGCGGGTCACAGAGTACCTTGTCTGTACACGCCACCCACGAGGGTGGT
CTGGAAATGCAGTTCGCTGTGCTGGGGACAGACGTCCACCATCATCCGGAGCTGGAG
CCTGGTGTGGAGTACTTTATCCGTGTATTTGCCATCCTGGAGAACAAGAGAGCATTCCT
GTCAGCGCCAGGGTGGCCACGTACTTACCTGCACCTGAAGGCCTGAAATTCAAGTCCATC
AAGGAGACATCTGTGGAAGTGGAGTGGATCCTCTAGACATTGCTTTTGAACCTGGGAG
ATCATCTTCCGGAATATGAATAAAGAAGATGAGGGAGAGATACCAAAAGCCTGAGGAGG
CCAGAGACCTCTTACCGGCAAACTGGTCTAGCTCCTGGGCAAGAGTATGAGATATCTCTG
CACATAGTGA AAAACAATACCCGGGGCCCTGGCCTGAAGAGGGTGACCACCACACGCTTG
GATGCCCCCAGCCAGATCGAGGTGAAGATGTCACAGACACCACTGCCTTGATCACCTGG
TTCAAGCCCCCTGGCTGAGATCGATGGCATTGAGCTGACCTACGGCATCAAAGACGTGCCA
GGAGACCGTACCACCATCGATCTCACAGAGGACGAGAACCAGTACTCCATCGGGAACCTG
AAGCCTGACACTGAGTACGAGGTGTCCCTCATCTCCCGCAGAGGTGACATGTCAAGCAAC
CCAGCCAAAGAGACCTTCAACACAGGCCTCGATGCTCCAGGAATCTTCGACGTGTTTCC
CAGACAGATAACAGCATCACCTTGAATGGAGGAATGGCAAGGCAGCTATTGACAGTTAC
AGAATTAAGTATGCCCCATCTCTGGAGGGGACCACGCTGAGGTTGATGTTCCAAGAGC
CAACAAGCCACAACCAAAACCACTCACAGGTCTGAGGCCGGGAATGAATATGGGATT
GGAGTTTCTGCTGTGAAGGAAGACAAGGAGAGCAATCCAGCGACCATCAACGCAGCCACA
GAGTTGGACACGCCCAAGGACCTTCAAGTTTCTGAACTGCAGAGACCAGCCTGACCTG
CTCTGGAAGACACCGTTGGCCAAATTTGACCGCTACCGCCTCAATTACAGTCTCCCCACA
GGCCAGTGGGTGGGAGTGCAGCTTCCAAGAAACACCACTTCTATGTCTGAGAGGCCTG
GAACCAGGACAGGAGTACATGTCTCTGACAGCCGAGAAAGGCAGACACAAGAGCAAG
CCCGCACGTGTGAAGGCATCCACTGAACAAGCCCTGAGCTGGAAAACCTCACCGTGACT
GAGGTTGGCTGGGATGGCCTCAGACTCAACTGGACCGCGGCTGACCAGGCCTATGAGCAC
TTTATCATTACAGGTGCAGGAGGCCAACAAGGTGGAGGCAGCTCGGAACCTCACCGTGCT
GGCAGCCTTCGGGCTGTGGACATACCGGGCCTCAAGGCTGCTACGCCTTATACAGTCTCC
ATCTATGGGGTGTCCAGGCTATAGAACACCAGTGCTCTCTGCTGAGGCCTCCACAGGG
GAAACTCCCAATTTGGGAGAGGTCGTGGTGGCCGAGGTGGGCTGGGATGCCCTCAAATC
AACTGGACTGCTCCAGAAGGGGCCTATGAGTACTTTTTCATTAGGTGCAGGAGGTGAC
ACAGTAGAGGCAGCCCAAGAACCTCACCGTCCCAGGAGGACTGAGGTCCACAGACCTGCCT
GGGCTCAAAGCAGCCACTCATTATACCATCACCATCCGCGGGGTCACTCAGGACTTCAGC
ACAACCCCTCTCTCTGTTGAAGTCTTGACAGAGGAGGTTCCAGATATGGGAAACCTCACA
GTGACCGAGGTTAGCTGGGATGCTCTCAGACTGAACTGGACCACGCCAGATGGAACCTAT
GACCAGTTTACTATTACAGTCCAGGAGGCTGACCAGGTGGAAGAGGCTCACAATCTCAG
GTTCTTGGCAGCCTGCGTTCCATGGAAATCCCAGGCCTCAGGGCTGGCCTCCTTACACA
GTCACCTTGACGGCGAGGTGAGGGGCCACAGCACTCGACCCCTTGCTGTAGAGGTGCTC
ACAGAGGATCTCCACAGCTGGGAGATTTAGCCGTGTCTGAGGTTGGCTGGGATGGCCTC
AGACTCAACTGGACCGCAGCTGACAAATGCCTATGAGCACTTGTCACTGAGGCTGAGGAG
GTCAACAAAGTGGAGGCAGCCAGAACCTCACGTTGCCTGGCAGCCTCAGGGCTGTGGAC
ATCCCGGGCCTCGAGGCTGCCACGCCTTATAGAGTCTCATCTATGGGGTGATCCGGGGC
TATAGAACACCAGTACTCTGCTGAGGCCTCCACAGCCAAAGAACCTGAAATTGGAAAC
TTAAATGTTTCTGACATAACTCCCGAGAGCTTCAATCTCTCTGGATGGCTACCGATGGG
ATCTTCGAGACCTTTACATTGAAATTTATTGATTCCAATAGGTTGCTGGAGCTGTGGAA
TATAATATCTCTGGTGTGAACGAACTGCCATATCTCAGGGCTACCCCTAGTACTGAT
TTTATTGTCTACCTCTCTGGACTGTCTCCAGCATCCGACCAAAACCATCAGTGCCACA

	<p>GCCACGACAGAGGCCCTGCCCTTCTGGAAAACCTAACCATTTCCGACATTAAATCCCTAC GGGTTACAGTTTCTCGATGGCATCGGAGAATGCCTTTGACAGCTTTCTAGTAACGGTG GTGGATTCTGGGAAGCTGCTGGACCCCAAGGAATTCACATTTTCAGGAACCCAGAGGAAG CTGGAGCTTAGAGGCCTCATAACTGGCATTGGCTATGAGGTTATGGTCTCTGGCTTCACC CAAGGGCATCAAACCAAGCCCTTGAGGGCTGAGATTGTTACAGAAGCCGAACCGGAAGTT GACAACCTTCTGGTTTCAGATGCCACCCAGACGGTTCCGTCTGTCTGGACAGCTGAT GAAGGGGTCTTCGACAAATTTTGTCTCAAATCAGAGATACAAAAAGCAGTCTGAGCCA CTGGAAATAACCTACTTGCCCCGAACGTACCAGGGACATAACAGGTCTCAGAGAGGCT ACTGAATACGAAATTGAACTCTATGGAATAAGCAAAGGAAGGCGATCCCAGACAGTCAGT GCTATAGCAACAACAGCCATGGGCTCCCCAAAGGAAGTCATTTTCTCAGACATCACTGAA AATTCGGCTACTGTCAGCTGGAGGGCACCACGGCCCAAGTGGAGAGCTTCGGATTACC TATGTGCCATTACAGGAGGTACACCTCCATGGTAACTGTGGACGGAACCAAGACTCAG ACCAGGCTGGTGAACCTCATACCTGGCGTGGAGTACCTTGTGAGCATCATCGCCATGAAG GGCTTTGAGGAAAGTGAACCTGTCTCAGGGTCATTACCACAGCTCTGGATGGCCCATCT GGCTTGGTGACAGCCAACATCACTGACTCAGAAGCCTTGCCAGGTGGCAGCCAGCCATT GCCACTGTGGACAGTTATGTCATCTCTACACAGGCGAGAAAGTGCCAGAAATTACACGC ACGGTGTCCGGGAACACAGTGGAGTATGCTCTGACCGACCTCGAGCCTGCCAGGAATAC ACACTGAGAATCTTTCAGAGAAAGGGCCCCAGAAGAGCTCAACCATCACTGCCAAGTTC ACAACAGACCTCGATTCTCAAGAGACTTGACTGCTACTGAGGTTAGTCGGAAACTGCC CTCCTTACCTGGCGACCCCCCGGGCATCAGTCACCGGTTACCTGCTGGTCTATGAATCA GTGGATGGCACAGTCAAGGAAGTCATTGTGGGTCCAGATACCACCTCCTACAGCCTGGCA GACCTGAGCCCATCCACCCACTACACAGCCAAGATCCAGGCACTCAATGGGCCCTGAGG AGCAATATGATCCAGACCATCTTACCACAATTGGACTCCTGTACCCCTTCCCCAAGGAC TGCTCCCAAGCAATGCTGAATGGAGACACGACCTCTGCGCTCTACACCATTATCTGAAT GGTGATAAGGCTCAGGCGCTGGAAGTCTTCTGTGACATGACCTCTGATGGGGGTGGATGG ATTGTGTTCTGAGACGCAAAACGGACGCGAGAACTTCTACAAAACCTGGAAGGCATAT GCTGCTGATTGTTGGGACCGCAGAGAAGAATTCTGGCTTGGGCTGGACAACCTGAACAAA ATCACAGCCAGGGGCGAGTACGAGCTCCGGGTGGACCTGCGGGACCATGGGAGACAGCC TTTGCTGTCTATGACAAGTTCAGCGTGGGAGATGCCAAGACTCGCTACAAGCTGAAGGTG GAGGGGTACAGTGGGACAGCAGGTGACTCCATGGCCTACCACAATGGCAGATCCTTCTCC ACCTTTGACAAGGACACAGATTACGCCATCACCAACTGTGCTCTGTCTACAAGGGGCTTC TGGTACAGGAACGTGTCACCGTGTCAACCTGATGGGGAGATATGGGGACAATAACACAGT CAGGGCGTTAACTGGTTCACCTGGAAGGGCCACGAACACTCAATCCAGTTTGTGAGATG AAGCTGAGACCAAGCAACTTCAGAAATCTTGAAGGCAGGCGAAACGGGCATAAATTGGA GGGACCACTGGGTGAGAGAGGAATAAGGCGGCCAGAGCGAGGAAGGATTTTACCAAAG CATCAATACAACCAGCCCAACCATCGGTCCACACCTGGGCATTGTTGGTGAATCAAAGCT GACCATGGATCCCTGGGGCCAACGGCAACAGCATGGGCCTCACCTCCTGTGATTTCCTT TCTTTGCACCAAGACATCAGTCTCCAACATGTTTCTGTTTTGTTGTTTATTGATTAGCAAA AATCTCCAGTGACAACATCGCAATAGTTTTTACTTCTCTTAGGTGGCTCTGGGATGGG AGAGGGGTAGGATGTACAGGGGTAGTTTTGTTTTAGAACCCAGCCGATTTTACATGAAGCT GTATAATTAATTGTCAATTATTTTGTAGCAAAGATTAAATGTGTCAATTGGAAGCCATCC CTTTTTTACATTTATACAACAGAAACAGAAAGCAATACTGTTTCCATTTTAAGGAT ATGATTAATATTATTAATATAATAATGATGATGATGATGATGATGATGATGATGATGATGAT AGAGATCTTTCTTTCCAAAACATTTCTGGACAGTACCTGATTGATTTTTTTTTTAAATA AAAGCACAAGTACTTTTGAACAAAAA</p>
	<p>ORF Start: ATG at 55</p>
	<p>ORF Stop: TAA at 6652</p>
	<p>SEQ ID NO: 182</p>
	<p>2199 aa</p>
	<p>MW at 240715.6kD</p>
<p>NOV35a, CG55832-01 Protein Sequence</p>	<p>MGAMTQLLAGVFLAFLALATEGGVLKKVIRHKRQSGVNATLPEENQPVVFNHVNLIKLPV GSQCSVDLESASGEKDLAPPSEPSSEFQEHVTDGENQIVFTHRINIPRRACGCAAAPDVK ELLSRLEELLENLVSSLREQCTAGAGCCLQPATGRLDTRPFCSGRGNFSTEGCGVCPEGW KGPNCSEPECNCHLRGRCDIDGQCICDDGFTGEDCSQLACPSDCNDQGKCVNGVCICFE GYAADCSRERICPVPCSEEHGTVDGLCVCHDGFAGDDCNKPLCLNNCYNRGRVENECVC DEGFTGEDCSELICPNDCFDRGRINGTCYCEEGFTGEDCGKPTCPHACHTQGRCEEQGC VCDEGFAGVDCSEKRCPADCHNRGRCDVGRCEDDGFTGADCGELKCPNGCSGHGRVCNG QCVDEGYTGEDCSQLRCPNDCHSRGRVVEGKCVCEQGFKGYDCSDMSPNDCHQHGRVCV NGMVCVDDGYTGEDCDRDRQCPDRCSNRGLCVDDGQVCEDGFTGPDCAELSCPNDCHGQGR CVNGQVCVCHGFMGKDCKEQRCPSDCHGQGRVVDGQCICHEGFTGLDCGQHSPPDCNNL GQCVSGRCICNEGYSGEDCSEVSPPKDLVVTEVTEETVNLAWDNEMRVTEYLVVYTPTHE</p>

	<p>GGLEMQFRVPGDQTSTII RELEPGVEYFIRVFAILENKKSIPVSARVATYLPAP EGLKFK SIKETSVEVEWDPLDIAFETWEIIFRNMNKEDEGEITKSLRRPETS YRQTGLAPGQEY EI SLHIVKNNTRGPGLKRVTTTRLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGIKD VPGDRTTIDLTEDENQYSIGNLKPDT EYEVSLISRRGDMSSNPAKETFTTGLDAPRNLR VSQTDNSITLEWRNGKAAIDSIRIKYAPISGGDHAEDVVPKSQQATTKTTLTGLRPGTEY GIGVSAVKEDKESNPATINAATELDTPKDLQVSETAETSLTLLWKTPLAKFDYRLNYSL PTGQWVGVLPRNTTSYVLRGLEPGQEYNVLLTAEKGRHKS KPARVKASTEQAPELENLT VTEVGWDGLRLNWTAAADQAYEHFI IQVQEANKVEAARNLTVPGSLRAVDIPGLKAATPYT VSIYGVIIQGYRTPVLSAEASTGETPNLGEVVVAEVGWDALKLNWTAPEGAYEYFFIIQVQE ADTVEAAQNLTVPGGRLSTDLPLGLKAATHYITIRGVTQDFSTTPLSVEVLTEEVPMGN LTVTEVSWDALRLNWTTPDGTQDFTIQVQEADQVEEAHNLTVPGSLRSMEIPLGRAGTP YTVTLHGEVRGHSTRPLAVEVVTEDLPQLGLD LAVSEVGWDGLRLNWTAAADNAYEHFVIQV QEVNKVEAAQNLTLPGLSLRAVDIPGLEAATPYRVSIYGVIRGYRTPVLSAEASTAKEPEI GNLNVSDITPESFNLSWMATDGI FETFTIEIIDS NRLETV EYNISGAERTAHISGLPPS TDFIVYLSGLAPSIRKTI SATATTEALPLENLTISDINPYGFTVSWMASENAFDSFLV TVVDSGKLLDPQEFTLSGTQRKLELRGLITGIGYEVMSGFTQGHQTKPLRAEIVTEAEP EVDNLLVSDATPDGFRLSWTADEGVDFDNFLKIRDTKKQSEPLEITLLAPERTRDITGLR EATEYEIELYGISKRRSQTVSAIATTAMGSPKEVIFSDITENSATVSWRAPTAQVESFR ITYVPITGGTSPMVTVDGTTQTRLVKLIPGVEYLVSI IAMKGFEESEPVSGSFTTALDG PSGLVTANITDSEALARWQPAIATVDSYVISYTG EKVPEITRTVSGNTVEYALTDL EAP EYTLRIFAEGFPQKSSITAKFTTDLDSPRDLTATEVQSETALLTWRPPRASVTGYLLVY ESVDGTVKEVIVGPDTSYSLADLSPSTHYTAKIQALNGPLRSNMIIQTIPTTIGLLYPPF KDCSQAMLNGDTSGLYTIYLNKDKAQALEVFCDMTSDGGGWIVFLRRKNGRENFYQNWK AYAAGFGDRREEFWLGLDNLNKITAQQQYELRVDLRDHGETAFAYVDKFSVGD AKTRYKL KVEGYSGTAGDSMAYHNRSFSTFDKDTDSAITNCALSTRGFWRNCHRVNLMGRYGDNN HSQGVNWFHWKGHEHSIQFAEMKLRPSNFRNLEGRRKRA</p>
	<p>SEQ ID NO: 183 7013 bp</p>
NOV35b, CG55832-03 DNA Sequence	<p>GAATTCGCTAGAGCCCTAGAGCCCCAGCAGCAGCCAGCCAAACCCACCTCCACCATGGGG GCCATGACTCAGCTGTTGGCAGGTGCTTTCTTGCTTTCCTTGCCTCGCTACCGAAGGT GGGGTCCTCAAGAAAGTCATCCGGCACAAGCGACAGAGTGGGGTGAACGCCACCCTGCCA GAAGAGAACCAGCCAGTGGTGTTTAAACACGTTTACAACATCAAGCTGCCAGTGGGATCC CAGTGTTCGGTGGATCTGGAGTCAGCCAGTGGGGAGAAAGACCTGGCACCCCTTCAGAG CCCAGCGAAAGCTTT CAGGAGCACACAGTAGATGGGGAAAACAGATTGTCTTACACAT CGCATCAACATCCCCCGCGGGCCTGTGGCTGTGCCGAGCCCTGATGTTAAGGAGCTG CTGAGCAGACTGGAGGAGCTGGAGAACCTGGTGTCTTCCCTGAGGGAGCAATGTACTGCA GGAGCAGGCTGTGTCTCCAGCCTGCCACAGGCCGCTTGGACACCAGGCCCTTCTGTAGC GGTCGGGGCAACTTCAGCACTGAAGGATGTGGCTGTGTCTGCGAACCTGGCTGGAAAGGC CCCAACTGCTCTGAGCCCGAATGTCCAGGCAACTGTACCTTCGAGGCCGCTGTCATTGAT GGGCAGTGCATCTGTGACGACGGCTTCAGGGCGAGGACTGCAGCCAGCTGGCTTGCCCC AGCGACTGCAATGACCAGGGCAAGTGCCTGAATGGAGTCTGCATCTGTTTCGAAGGCTAC CGGGCTGACTGCAGCCGTGAATCTGCCAGTGCCCTGCAGTGAGGAGCACGGCACATGT GTAGATGGCTTGTGTGTGTGCCAGATGGCTTTCAGGCGATGACTGCAACAAGCCTCTG TGTCTCAACAATTGTCTACAACCGTGGACGATGCGTGGAGAATGAGTGCCTGTGTGATGAG GGTTTCAGGGCGAAGACTGCAGTGAGCTCATCTGCCCAATGACTGCTTCGACCGGGGC CGCTGCATCAATGGCACCTGCTACTGCGAAGAAGGCTTACAGGTGAAGACTGCGGGAAA CCCACCTGCCACATGCCTGCCACACCCAGGGCCGGTGTGAGGAGGGGCAGTGTGTATGT GATGAGGGCTTTGCCGGTGTGGACTGCAGCGAGAAGAGGTGCTCTGCTGACTGTCACAAT CGTGGCCGCTGTGTAGACGGCGGTGTGAGTGTGATGATGGTTTCACTGGAGCTGACTGT GGGGAGCTCAAGTGTCCCAATGGCTGCAGTGGCCATGGCCGCTGTGTCAATGGGCAGTGT GTGTGTGATGAGGGCTATCTGGGGAGGACTGCAGCCAGCTACGGTGCCTCAATGACTGT CACAGTCGGGGCCGCTGTGTGAGGGCAAATGTGTATGTGAGCAAGGCTTCAAGGGCTAT GACTGCAGTGACATGAGCTGCCCTAATGACTGTCAACAGCACGGCCGCTGTGTGAATGGC ATGTGTGTTTGTGATGACGGCTACACAGGGGAAGACTGCCGGGATCGCCCAATGCCCCAGG GACTGCAGCAACAGGGGCCTCTGTGTGGACGGACAGTGGCTGTGTGAGGACGGCTTACC GGCCCTGACTGTGCAGAACTCTCTGTCCAAATGACTGCCATGGCCAGGGTCTGTGTG AATGGGCAGTGCCTGTGCCATGAAGGATTATGGGCAAGACTGCAAGGAGCAAGATGT CCCAGTGACTGTATGGCCAGGGCCGCTGCGTGGACGGCCAGTGCATCTGCCACGAGGGC TTCACAGGCCTGGACTGTGGCCAGCACTCTGCCCCAGTGACTGCAACAACCTAGGACAA TGCGTCTCGGGCCGCTGCATCTGCAACGAGGGCTACAGCGGAGAAGACTGCTCAGAGGTG</p>

TCTCCTCCCAAAGACCTCGTTGTGACAGAAGTGACGGAAGAGACGGTCAACCTGGCCTGG
GACAATGAGATGCGGGTCACAGAGTACCTTGTCTGTACACGCCACCCACGAGGGTGGT
CTGGAATGCAGTTCCGTGTGCTGGGGACCAGACGTCCACCATCATCCGGGAGCTGGAG
CCTGGTGTGGAGTACTTTATCCGTGTATTGCCATCCTGGAGAACAGAAGAGCATTCCCT
GTCAGCGCCAGGGTGGCCACGTACTTACCTGCACCTGAAGGCCTGAAATTCAAGTCCATC
AAGGAGACATCTGTGGAAGTGGAGTGGGATCCTCTAGACATTGCTTTTGAACCTGGGAG
ATCATCTTCCGGAATATGAATAAAGAAGATGAGGGAGAGATCACCAAAAGCCTGAGGAGG
CCAGAGACCTCTTACCGGCAACTGGTCTAGCTCCTGGGCAAGAGTATGAGATATCTCTG
CACATAGTGAAAAACAATACCCGGGGCCCTGGCCTGAAGAGGGTGACCACCACACGCTTG
GATGCCCCCAGCCAGATCGAGGTGAAAGATGTCACAGACACCACTGCCTTGATCACCTGG
TTCAAGCCCCCTGGCTGAGATCGATGGCATTGAGCTGACCTACGGCATCAAAGACGTGCCA
GGAGACCGTACCACCATCGATCTCAGAGGACGAGAACCAGTACTCCATCGGGAACCTG
AAGCCTGACACTGAGTACGAGGTGTCCCTCATCTCCCGCAGAGGTGACATGTCAAGCAAC
CCAGCCAAAGAGACCTTCACAACAGGCCTCGATGCTCCAGGAATCTTCGACGTGTTTCC
CAGACAGATAACAGCATCACCTTGAATGGAGGAATGGCAAGGCAGCTATTGACAGTTAC
AGAATTAAGTATGCCCCATCTCTGGAGGGGACCACGCTGAGGTGATGTTCCAAAGAGC
CAACAAGCCACAACCAAAACCACTCACAGGTCTGAGGCCGGGAATGAATATGGGATT
GGAGTTTCTGCTGTGAAGGAAGACAAGGAGAGCAATCCAGCGACCATCAACGCAGCCACA
GAGTTGGACACGCCCAAGGACCTTCAGGTTTCTGAAACTGCAGAGACCAGCCTGACCTG
CTCTGGAAGACACCGTTGGCCAAATTTGACCGCTACCGCCTCAATTACAGTCTCCCCACA
GGCCAGTGGGTGGGAGTGCAGCTTCCAAGAAACACCACTTCTATGTCTGAGAGGCCTG
GAACCAGGACAGGAGTACAATGTCTCTCTGACAGCCGAGAAAGGCAGACACAAGAGCAAG
CCCGCACGTGTGAAGGCATCCACTGAACAAGCCCCTGAGCTGGAAAACCTCACCGTGACT
GAGGTTGGCTGGGATGGCCTCAGACTCAACTGGACCGCGGCTGACCAGGCCTATGAGCAC
TTTATCATTCAGGTGACGAGGCCAACAGGTGGAGGCAGCTCGGAACCTCACCGTGCCT
GGCAGCCTTCGGGCTGTGGACATACCGGGCCTCAAGGCTGCTACGCCTTATACAGTCTCC
ATCTATGGGGTGATCCAGGGCTATAGAACACCAAGTCTCTCTGCTGAGGCCTCCACAGGG
GAAACTCCCAATTTGGGAGAGGTCTGTGGTGGCCGAGGTGGGCTGGGATGCCCTCAAACCTC
AACTGGACTGCTCCAGAAGGGGCCTATGAGTACTTTTTCATTACAGGTGCAGGAGGCTGAC
ACAGTAGAGGCAGCCAGAACCTCACCGTCCCAGGAGGACTGAGGTCCACAGACCTGCCT
GGGCTCAAAGCAGCCACTCATTATACCATCACCATCCGCGGGGTCACTCAGGACTTCAGC
ACAACCCCTCTCTCTGTTGAAGTCTTGACAGAGGAGGTTCCAGATATGGGAAACCTCACA
GTGACCAGGTTAGCTGGGATGCTCTCAGACTGAACTGGACCACGCCAGATGGAACCTAT
GACCAGTTTACTATTACAGTCCAGGAGGCTGACCAGGTGGAAGAGGCTCACAATCTCACG
GTTCTCTGGCAGCCTGCGTTCCATGGAATCCCAGGCCTCAGGCTGGCACTCCTTACACA
GTCACCTGCACGGCAGGTGAGGGGCCACAGCACTCGACCCCTTGCTGTAGAGGTGCTC
ACAGAGGATCTCCACAGCTGGGAGATTAGCCGTGTCTGAGGTGGCTGGGATGGCCTC
AGACTCAACTGGACCGCAGCTGACAATGCCTATGAGCACTTTGTCAATTACAGGTGCAGGAG
GTCAACAAAGTGGAGGCAGCCAGAACCTCACGTTGCCTGGCAGCCTCAGGCTGTGGAC
ATCCCGGGCCTCGAGGCTGCCACGCCTTATAGAGTCTCCATCTATGGGGTGATCCGGGGC
TATAGAACACCACTACTCTCTGCTGAGGCCTCCACAGCCAAAGAACCTGAAATTGGAAC
TTAAATGTTTCTGACATAACTCCCGAGAGCTTCAATCTCTCTGGATGGCTACCGATGGG
ATCTTCGAGACCTTACCATTGAAATTTATTGATTCCAATAGGTTGCTGGAGACTGTGGAA
TATAATATCTCTGGTGCTGAACGAACTGCCCATATCTCAGGGCTACCCCTAGTACTGAT
TTTATTGTCTACCTCTCTGGAATTGCTCCAGCATCCGGAACCAACCATCAGTGCCACA
GCCACGACAGAAGCCGAACCGGAAGTTGACAACCTTCTGGTTTCAGATGCCACCCAGAC
GGTTTCCGTCTGTCTTGACAGCTGATGAAGGGGTCTTCGACAATTTTGTCTCAAATC
AGAGATACCAAAAAGCAGTCTGAGCCACTGGAATAACCTACTTGCCCCCGAACGTACC
AGGGACATAACAGGTCTCAGAGAGGCTACTGAATACGAAATTGAACTCTATGGAATAAGC
AAAGGAAGGCGATCCAGACAGTCAAGTCTATAGCAACAACAGCCATGGGCTCCCCAAAG
GAAGTCATTTTCTCAGACATCACTGAAAATTCGGCTACTGTCAAGCTGGAGGGCACCCACG
GCCAAGTGGAGAGCTTCCGGATTACCTATGTGCCATTACAGGAGGTACACCTCCATG
GTAACCTGTGGACGGAACCAAGACTCAGACCAGGCTGGTGAAACTCAGAGAGAAAGGCCGAG
TACCTTGTCAGCATCATCGCCATGAAGGGCTTTGAGGAAAGTGAACCTGTCTCAGGGTCA
TTCACCACAGCTCTGGATGGCCCATCTGGCCTGGTGACAGCCAACATCACTGACTCAGAA
GCCTTGCCAGGTGGCAGCCAGCCATTGCCACTGTGGACAGTTATGTACTCTCTACACA
GGCGAGAAAGTGCAGAAATTACACGCACGGTGTCCGGGAACACAGTGGAGTATGCTCTG
ACCGACCTCGAGCCTGCCAGGAATACACACTGAGAATCTTTGAGAGAGAAAGGCCCCAG
AAGAGCTCAACCATCACTGCAAGTTTCAACAGACCTCGATTCTCAAAGAGACTTGACT
GCTACTGAGGTTCACTCGGAAACTGCCCTCCTTACCTGGCGACCCCCCGGCATCAGTC

	<p>ACCGGTTACCTGCTGGTCTATGAATCAGTGGATGGCACAGTCAAGGAAGTCATTGTGGGT CCAGATACCACCTCTACAGCCTGGCAGACCTGAGCCCATCCACCCACTACACAGCCAAG ATCCAGGCACTCAATGGGCCCCCTGAGGAGCAATATGATCCAGACCATCTTCACCACAATT GGACTCCTGTACCCCTTCCCCAAGGACTGCTCCCAAGCAATGCTGAATGGAGACACGACC TCTGGCCTCTACACCATTATCTGAATGGTGATAAGGCTCAGGCGCTGGAAGTCTTCTGT GACATGACCTCTGATGGGGTGGATGGATTGTGTTCTGAGACGCAAAAACGGACGCGAG AACTTCTACCAAACTGGAAGGCATATGCTGCTGGATTGGGGACCGCAGAGAAGAATTC TGGCTTGGGCTGGACAACCTGAACAAAATCACAGCCAGGGGAGTACGAGCTCCGGGTG GACCTGCGGACCATTGGGGAGACAGCCTTGTCTGTATGACAGTTTACGCGTGGGAGAT GCCAAGACTCGCTACAAGCTGAAGGTGGAGGGGTACAGTGGGACAGCAGGTGACTCCATG GCCTACCACAATGGCAGATCCTTCTCCACCTTTGACAAGGACACAGATTACGCCATCACC AACTGTGCTCTGTCTACAAGGGCTTCTGGTACAGGAAGTGTACCGTGTCAACCTGATG GGGAGATATGGGGACAATAACACAGTCAGGGCGTTAACTGGTTCCACTGGAAGGGCCAC GAACACTCAATCCAGTTTGTCTGAGATGAAGCTGAGACCAAGCAACTTCAGAAATCTTGAA GGCAGGCGCAACGGGCATAAATTGGAGGGACCACTGGGTGAGAGAGGAATAAGGCGGCC CAGAGCGAGGAAAGGATTTTACCAAAGCATCAATACAACCAGCCCAACCATCGGTCCACA CCTGGGCATTGGTGAGAATCAAAGCTGACCATGGATCCCTGGGGCCAACGGCAACAGCA TGGGCCTCACCTCCTCTGTGATTTCTTTCTTGCACCAAGACATCAGTCTCCAACATGT TTCTGTTTGTGTTGATTGATTGAGCAAAAATCTCCAGTGACAACATCGCAATAGTTTTTT ACTTCTCTAGGTGGCTCTGGGATGGGAGAGGGGTAGGATGTACAGGGGTAGTTTGT AGAACCAGCCGTATTTACATGAAGCTGTATAATTAAATTGTCATTATTTTGTAGCAAA GATTAAATGTGTCTATTGGAAGCCATCCCTTTTTTACATTTATACAACAGAAACCAGAA AAGCAATACTGTTTCCATTTTAAGGATATGATTAAATATTATATAATAATGATGATG ATGATGATGAAACTAAGGATTTTCAAGAGATCTTCTTTCCAAAACATTCTGGACAG TACCTGATTGTATTTTTTTTTTAAATAAAAGCACAAAGTACTTTTGAAAAAAA</p>
	<p>ORF Start: ATG at 55</p>
	<p>ORF Stop: TAA at 6379</p>
	<p>SEQ ID NO: 184</p>
	<p>2108 aa</p>
	<p>MW at 230729.3kD</p>
<p>NOV35b, CG55832-03 Protein Sequence</p>	<p>MGAMTQLLAGVFLAFLALATEGGVLKKVIRHKRQSGVNATLPEENQPVVFNHVYNIKLPV GSQCSVDLESASGEKDLAPPSESPESFQEHTEVDGENQIVFTHRINIPRRACGCAAPDVK ELLSRLEELNVLSSLRQCTAGAGCCLQPATGRLDTRPFCSGRGNFSTEGCGVCPEGW KGPNCSEPECPCGNCHLRGRICDQCICDDGFTGEDCSQLACPSDCNDQKGCNVGVCIFE GYAADCSREICPVPCSEEHGTCVDGLCVCHDGFAGDDCNKPLCLNNCYNRGRVCVENECVC DEGFTGEDCSELICPNDCFDGRRCINGTCYCEEGFTGEDCGKPTCPHACHTQGRCEEQGC VCDEGFAGVDCSEKRCPCADCHNRGRCDVGRCECDDGFTGADCGELKCPNGCSGHGRVCNG QCVCEGYTGEDCSQLRCPNDCHSRGRVCVEGKVCCEQGFKG YDCSDMSPNDCHQHGRCV NGMVCDDGYTGEDCDRDRQCPDRCNSRGLCVDGQVCEDGFTGPDCAELSCPNDCHGQGR CVNGQVCVHEGFMGKDCKEQRCPSDCHGQGRCDVGCICHEGFTGLDCGQHSCPSDCNNL GQCVSGRCICNEGYSGEDCSEVSPPKDLVVTEVTEETVNLAWDNEMRVTELVVYTPTHE GGLEMQFRVPGDQSTIIIRELEPGVEYFIRVFAILNKKSI PVSARVATYLPAPGLKFK SIKETSVEVEWDPDLIAFETWEII FRNMNKEDEGEITKSLRRPETS YRQTGLAPGQEYEI SLHIVKNNTRGPGLKRVTTTRLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGKD VPGDRTTIDLTEDENQYSIGNLKPDEYEVSLISRRGDMSSNPAKETFTTGLDAPRNLRR VSQTDNSITLEWRNGKAAIDSYRIKYAPISGGDHAEVDVPKQQATTKTTLTGLRPGTEY GIGVSAVKEDKESNPATINAATELDTPKDLQVSETAETSLTLLWKTPLAKFDRYRLNYSL PTGQWVGVLPRNTTSYVLRGLEPGQEYNVLLTAEKGRHKS KPARVKASTEQAPELENLT VTEVGWDGLRLNWTAAADQAYEHFI IQVQEANKVEAARNLTVPGLRAVDIPGLKAATPYT VSIYGVIIQGYRTPVLSAEASTGETPNLGEVVVAEVGWDALKLNWTAPEGAYEYFFIIQVQE ADTVEAAQNLTVPGLRSTDLPLGLKAATHYITIRGVTQDFSTPLSVEVLTEVPDMGN LTVTEVSWDALRLNWTTPDGYDQFTIQVQEADQVEEAHNLTPGSLRSMEIPGLRAGTP YTVTLHGEVRGHSTRPLAVEVVTEDLPQLGDLAVSEVGWDGLRLNWTAAADNAYEHFVIQV QEVNKVEAAQNLTLPGLRAVDIPGLEAATPYRVSIYGVIRGYRTPVLSAEASTAKEPEI GNLNVSDITPESFNLSWMATDGFETFTIEIIDSNRLLTVEYNI SGAERTAHISGLPPS TDFIVYLSGLAPSIRTKTISATATTEAEFEVDNLLVSDATPDGFRLSWTADEGVFDNFVL KIRDTKKQSEPLEITLLAPERTRDITGLREATEYEIELYGISKGRRSQTVSAIATTAMGS PKEVIFSDITENSATVSWRAPTAQVESFRITYVPITGGTPSMVTVDGTKTQTRLVKLIPG VEYLVSI IAMKFEESEPVSFTTALDGPSTGLVTANITDSEALARWQPAIATVDYSVIS YTGEKVPEITRTVSGNTVEYALTDLEPATEYTLRI FAEKGPQKSSITITAKFTTDLDSPRD LTATEVQSETALLTWRPPRASVTGYLLVYESVDGTVKEVIVGPDTSYSLADLSPSTHYT</p>

	AKIQALNGLPLRSNMIQTIFFTIGLLYPPFKDCSQAMLNDDTSSGLYTIYLNKDKAQALEV FCDMTSDGGGWIVFLRRKNGRENFYQNWKAYAAGFGDRREEFWLGLDNLNKITAQGGYEL RVDLRDHGETAFAVYDKFSVGDAKTRYKLVKVEGYSGTAGDSMAYHNGRSFSTFDKDTDSA ITNCALSTRGFWYRNCHRVNLMGRYGDNNHSQGVNWFHWKGHEHSIQFAEMKLRLPSNFRN LEGRRKRA		
	SEQ ID NO: 185	5375 bp	
NOV35c, CG55832-02 DNA Sequence	GAATTCGCTAGAGCCCTAGAGCCCCAGCAGCACCCAGCCAAACCCACCTCCACCATGGGG GCCATGACTCAGCTGTTGGCAGGTGTCTTCTTGCTTTCCTTGCCCTCGCTACCGAAGGT GGGGTCTCTCAAGAAAGTCATCCGGCACAAGCGACAGAGTGGGGTGAACGCCACCTGCCA GAAGAGAACCAGCCAGTGGTGTTTAACCACGTTTACAACATCAAGCTGCCAGTGGGATCC CAGTGTTCGGTGGATCTGGAGTCAGCCAGTGGGGAGAAAGACCTGGCACCGCCTTCAGAG CCCAGCGAAAGCTTTCAGGAGCACACAGTAGATGGGGAAAACAGATTGTCTTACACAT CGCATCAACATCCCCCGCCGGGCTGTGGCTGTGCCGAGCCCTGATGTTAAGGAGCTG CTGAGCAGACTGGAGGAGCTGGAGAACCTGGTGTCTTCCCTGAGGGAGCAATGTAAGTCA GGAGCAGGCTGTGTCTCCAGCCTGCCACAGGCCGCTTGACACCAGGCCCTTCTGTAGC GGTCGGGGCAACTTCAGCACTGAAGGATGTGGCTGTGTCTGCGAACCTGGCTGAAAAGGC CCCAACTGCTCTGAGCCGAATGTCCAGGCAACTGTACCTTCGAGGCCGGTGCATTGAT GGGCAGTGCATCTGTGACGACGGCTTCACGGGCGAGGACTGCAGCCAGTGGCTTGGCTTGC AGCGACTGCAATGACCAGGGCAAGTGCCTGAATGGAGTCTGCATCTGTTTCGAAGGCTAC GCGGCTGACTGCAGCCGTGAAATCTGCCAGTGCCCTGCAGTGAGGAGCACGGCACATGT GTAGATGGCTTGTGTGTGTGCCACGATGGCTTTCAGGCGATGACTGCAACAAGCCTCTG TGTCTCAACAATTGCTACAACCGTGGACGATGCGTGGAGAATGAGTGCGTGTGTGATGAG GGTTTCACGGGCGAAGACTGCAGTGAGCTCATCTGCCCAATGACTGCTTCGACCGGGGC CGCTGCATCAATGGCACCTGCTACTGCGAAGAAGGCTTCACAGGTGAAGACTGCGGGAAA CCACCTGCCACATGCCTGCCACACCCAGGGCCGGTGTGAGGAGGGGCAGTGTGTATGT GATGAGGGCTTTGCGGGTGTGGACTGCAGCGAGAAGAGGTGTCTGCTGACTGTCACAAT CGTGGCCGCTGTGTAGACGGGCGGTGTGAGTGTGATGATGGTTTCACTGGAGCTGACTGT GGGGAGCTCAAGTGTCCCAATGGCTGCAGTGGCCATGGCCGCTGTGTCAATGGGCAGTGT GTGTGTGATGAGGGCTATACTGGGGAGGACTGCAGCCAGCTACGGTGCCCAATGACTGT CACAGTCGGGGCCGCTGTGTGAGGGCAATGTGTATGTGAGCAAGGCTTCAAGGGCTAT GACTGCAGTGACATGAGCTGCCCTAATGACTGTACCAGCACGGCCGCTGTGTGAATGGC ATGTGTGTTTGTGATGACGGCTACACAGGGGAAGACTGCCGGGATCGCCAATGCCCCAGG GACTGCAGCAACAGGGGCTCTGTGTGGACGGACAGTGCCTCTGTGAGGACGGCTTCACC GGCCCTGACTGTGCAGAACTCTCTGTCCAAATGACTGCCATGGCCAGGGTCTGTGTGTG AATGGGCAGTGCCTGTGCCATGAAGATTATGGGCAAGACTGCAAGGAGCAAGATGT CCCAGTACTGTATGGCCAGGGCCGCTGCGTGGACGGCCAGTGCATCTGCCACGAGGGC TTCACAGGCCTGGACTGTGGCCAGCACTCTGCCCAAGTGAAGTCAACAATAGGACAA TGCGTCTCGGGCCGCTGCATCTGCAACGAGGGCTACAGCGGAGAAGACTGCTCAGAGGTG TCTCTCCCAAAGACCTCGTTGTGACAGAAGTGACGGAAGAGACGGTCAACCTGGCCTGG GACAATGAGATGCGGGTACAGAGTACCTTGTCTGTACACGCCACCCACGAGGGTGGT CTGGAATGACAGTTCCTGTGCTGGGACCAGACGTCCACCATCATCCGGGAGCTGGAG CCTGGTGTGGAGTACTTTATCCGTGTATTGCCATCTTGAGAAACAAGAAGAGCATTCCCT GTCAGCGCCAGGGTGGCCACGTACTTACCTGCACCTGAAGGCCGTGAAATCAAGTCCATC AAGGAGACATCTGTGGAAGTGGAGTGGGATCCTCTAGACATTGCTTTTGAAACCTGGGAG ATCATCTTCCGGAATATGAATAAAGAAGATGAGGGAGAGATCACCAAAAGCCTGAGGAGG CCAGAGACCTCTTACCGGCAAACTGGTCTAGCTCCTGGGCAAGAGTATGAGATATCTCTG CACATAGTGA AAAACAATACCCGGGGCCCTGGCCTGAAGAGGGTGACCACCACACGCTTG GATGCCCCCAGCCAGATCGAGGTGAAAGATGTACAGACACCACTGCCTTGATCACCTGG TTCAAGCCCTGGCTGAGATCGATGGCATTGAGCTGACCTACGGCATCAAAGACGTGCCA GGAGACCGTACCACCATCGATCTCACAGAGGACGAGAACCAGTACTCCATCGGGAACCTG AAGCCTGACACTGAGTACGAGGTGTCCCTCATCTCCCGCAGAGGTGACATGTCAAGCAAC CCAGCCAAAGAGACCTTCACAACAGGCCTCGATGCTCCAGGAATCTTCGACGTGTTTCC CAGACAGATAACAGCATACCTTGAATGGAGGAATGGCAAGGCAGCTATTGACAGTTAC AGAATTAAGTATGCCCCATCTCTGGAGGGGACCAGCTGAGGTTGATGTTCCAAAGAGC CAACAAGCCACAACCAAAACCACTCACAGGTCTGAGGCCGGGAATGAATATGGGATT GGAGTTTCTGCTGTGAAGGAAGACAAGGAGAGCAATCCAGCGACCATCAACGCAGCCACA GAGTTGGACACGCCCAAGGACCTTCAGGTTTCTGAACTGCAGAGACCAAGCCTGACCCTG CTCTGGAAGACACCGTTGGCCAAATTGACCGCTACCGCCTCAATTACAGTCTCCCCACA GGCCAGTGGGTGGGAGTGCAGCTTCAAGAAACACCACTTCCTATGTCTGAGAGGCTG		

	GAACCAGGACAGGAGTACAATGTCTCTCTGACAGCCGAGAAAGGCAGACACAAGAGCAAG CCCGCACGTGTGAAGGCATCCACTGCCATGGGCTCCCCAAAGGAAGTCATTTTCTCAGAC ATCACTGAAAAATTCGGCTACTGTCTGAGCTGGAGGGCACCCACAGCCCAAGTGGAGAGCTTC CGGATTACCTATGTGCCATTACAGGAGGTACACCCCTCCATGGTAACTGTGGACGGAACC AAGACTCAGACCAGGCTGGTGAACCTCATACCTGGCGTGGAGTACCTTGTCTCAGCATCATC GCCATGAAGGGCTTTGAGGAAAGTGAACCTGTCTCAGGGTCATTACCCACAGCTCTGGAT GGCCCATCTGGCCTGGTGACAGCCAACATCACTGACTCAGAAGCCTTGGCCAGGTGGCAG CCAGCCATTGCCACTGTGGACAGTTATGTCATCTCTTACACAGGCAGAAAGTGCCAGAA ATTACACGCACGGTGTCCGGGAACACAGTGGAGTATGCTCTGACCGACCTCGAGCCTGCC ACGGAATACACACTGAGAATCTTTGCAGAGAAAGGGCCCCAGAAGAGCTCAACCATCACT GCCAAGTTCACAACAGACCTCGATTCTCCAAGAGACTTGACTGCTACTGAGGTTTCAGTCG GAAACTGCCCTCCTTACCTGGCGACCCCCCGGGCATCAGTCACCGGTTACCTGCTGGTC TATGAATCAGTGGATGGCACAGTCAAGGAAGTCATTGTGGGTCCAGATACCACCTCCTAC AGCCTGGCAGACCTGAGCCCATCCACCCACTACACAGCCAAGATCCAGGCACTCAATGGG CCCCTGAGGAGCAATATGATCCAGACCATCTTCACCACAATTGGACTCTGTACCCCTTC CCCAAGGACTGCTCCCAAGCAATGCTGAATGGAGACACGACCTCTGGCCTCTACACCATT TATCTGAATGGTGATAAGGCTCAGGCGCTGGAAGTCTTCTGTGACATGACCTCTGATGGG GGTGGATGGATTGTGTTCTGAGACGCAAAAACGGACGCGAGAACTTCTACCAAACTGG AAGGCATATGCTGTGATTGTGGGGACCGCAGAGAAGAACTCTGGCTTGGGCTGGACAAC CTGAACAAAATCAGAGCCAGGGGCAGTACGAGCTCCGGGTGGACCTGCGGGACCATGGG GAGACAGCCTTTGTCTGTCTATGACAAGTTCAGCGTGGGAGATGCCAAGACTCGCTACAAG CTGAAGGTGGAGGGGTACAGTGGGACAGCAGGTGACTCCATGGCCTACCACAATGGCAGA TCCTTCTCCACCTTTGACAAGGACACAGATTCAAGCATCACCACCTGTGCTCTGTCTACA AGGGGCTTCTGGTACAGGAAGTGTCAACCTGATGGGGAGATATGGGGACAAT AACCACAGTCAGGGCGTTAACTGGTTCCTGGAAGGGCCACGAACACTCAATCCAGTTT GCTGAGATGAAGCTGAGACCAAGCAACTTCAGAAATCTGAAGGCAGGCGCAACCGGCA TAAATTGGAGGGACCACTGGGTGAGAGAGGAATAAGCGGCCAGAGCGAGGAAAGGATT TTACCAAGCATCAATACAACAGCCCCAACCATCGGTCCACACCTGGGCATTGTGGTGAGA ATCAAAGCTGACCATGGATCCCTGGGGCCAACGGCAACAGCATGGGCCTCACCTCCTCTG TGATTTCTTTCTTTGCACCAAGACATCAGTCTCCAACATGTTTCTGTTTTGTGTTTGA TTCAGCAAAAATCTCCAGTGACAACATCGCAATAGTTTTTTACTTCTCTCTAGGTGGCTC TGGGATGGGAGAGGGGTAGGATGTACAGGGGTAGTTTGTGTTTAGAACCAGCGGTATTTTA CATGAAGCTGTATAATTAATTGTCTATTATTTTGTGTTAGCAAAGATTAAATGTGTCTATTGG AAGCCATCCCTTTTTTACATTTTATACAACAGAAACAGAAAAGCAATATGTTTCCAT TTTAAGGATATGATTAATATTATTAATAATAATGATGATGATGATGATGAAAACTAAG GATTTTTCAAGAGATCTTTCTTCCAAAACATTTCTGGACAGTACCTGATTGTATTTTTT TTTTAAATAAAAGCACAGTACTTTTGAAAAAAA		
	ORF Start: ATG at 55		ORF Stop: TAA at 4741
	SEQ ID NO: 186	1562 aa	MW at 171222.6kD
NOV35c, CG55832-02 Protein Sequence	MGAMTQLLAGVFLAFLALATEGGVLKKVIRHKRQSGVNATLPEENQPVVFNHVNLIKLPV GSQCSVDLESASGEKDLAPPSEPSSEFQHTVDGENQIVFTHRINIPRRACGCAAAPDVK ELLSRLEELLENLVSSLREQCTAGAGCCLQPATGRDLTRPFCSGRGNFSTEGCGVCEPGW KGPNCSEPECNCHLRGRCIDGQCICDDGFTGEDCSQLACPSDCNDQGKCVNGVCICFE GYAADCSTREICPVPCSEEHGTCVDGLCVCHDGFAGDDCNKPLCLNLCYNRGRVENECVC DEGFTGEDCSELICPNDCFDGRGRINGTCYCEEGFTGEDCGKPTCPHACHTQGRCEEQGC VCDEGFAGVDCSEKRCPADCHNRGRCDVGRCEDDGFTGADCGELKCPNGCSGHGRVCNG QCVCEGYTGEDCSQLRCPNDCHSRGRVCEGKCVCEQGFKGYDCSDMSPNDCHQHGRVCV NGMCVDDGYTGEDCRDRQCPRDCSNRGLCVDGQCVCEDGFTGPDCAELSPNDCHGQGR CVNGQCVCHGFMGKDCKEQRCPDCHGQGRCDVGGQICHEGFTGLDCGQHSCPSDCNNL GQCVSGRCICNEGYSGEDCSEVSPPKDLVVTEVTEETVNLAWDNEMRVTEYLVVYTPTHE GGLEMQRVPGDQTSTIIRELEPGVEYFIRVFAILENKKSIPIVSARVATLPAPEGLKFK SIKETSVEVEWDPLDIAFETWEIIFRNMNKEDEGEITKSLRRPETSRYQTGLAPGEYEI SLHIVKNNTGRPGLRVTTTRLDAPSQIEVKDVTDTTALITWFKPLAEIDGIELTYGIKD VPGDRTTIDLTEDENQYSIGNLKPDTYEVSLSIRRGDMSSNPAKETFTTGLDAPRNLRR VSQTDNSITLWRNGKAAIDSYRIKYAPISGGDHAEVDVPSQQAATTKTTLTGLRPGTEY GIGVSAVKEDKESNPATINAATELTPKDLQVSETAETSLTLLWKTPLAKFDTRYRLNYSL PTGQWVGVLPRNTTSYVLRGLEPGQEYNVLLTAEKGRHKSPPARVKAATAMGSPKEVIF SDITENSATVSWRAPTAQVESFRITYVPITGGTPSMVTVDGTKTQTRLVKLIPGEYVLVS		

I I A M K G F E E S E P V S G S F T T A L D G P S G L V T A N I T D S E A L A R W Q P A I A T V D S Y V I S Y T G E K V P E I T R T V S G N T V E Y A L T D L E P A T E Y T L R I F A E K G P Q K S S T I T A K F T T D L D S P R D L T A T E V Q S E T A L L T W R P P R A S V T G Y L L V Y E S V D G T V K E V I V G P D T T S Y S L A D L S P S T H Y T A K I Q A L N G P L R S N M I Q T I F T T I G L L Y P F P K D C S Q A M L N G D T T S G L Y T I Y L N G D K A Q A L E V F C D M T S D G G G W I V F L R R K N G R E N F Y Q N W K A Y A A G F G D R R E E F W L G L D N L N K I T A Q G Q Y E L R V D L R D H G E T A F A V Y D K F S V G D A K T R Y K L K V E G Y S G T A G D S M A Y H N G R S F S T F D K D T D S A I T N C A L S T R G F W Y R N C H R V N L M G R Y G D N N H S Q G V N W F H W K G H E H S I Q F A E M K L R P S N F R N L E G R R K R A
--

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 35B.

Table 35B. Comparison of NOV35a against NOV35b and NOV35c.		
Protein Sequence	NOV35a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV35b	1..1884 1..1881	1595/1886 (84%) 1660/1886 (87%)
NOV35c	1..1332 1..1327	1079/1335 (80%) 1120/1335 (83%)

Twelve polymorphic variants of NOV35c have been identified and are shown in Table 41N.

- 5 Further analysis of the NOV35a protein yielded the following properties shown in Table 35C.

Table 35C. Protein Sequence Properties NOV35a	
PSort analysis:	0.8200 probability located in endoplasmic reticulum (membrane); 0.1900 probability located in plasma membrane; 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in outside
SignalP analysis:	Cleavage site between residues 23 and 24

A search of the NOV35a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 35D.

10

Table 35D. Geneseq Results for NOV35a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV35a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value

AAR94562	Human cytotactin - <i>Homo sapiens</i> , 2199 aa. [WO9608513-A1, 21-MAR-1996]	1..2199 1..2199	2199/2199 (100%) 2199/2199 (100%)	0.0
AAB36935	Human tenascin-C - <i>Homo sapiens</i> , 2201 aa. [WO200066628-A1, 09-NOV-2000]	1..2199 1..2201	2194/2201 (99%) 2198/2201 (99%)	0.0
AAR94563	Chicken cytotactin - <i>Gallus</i> sp, 1810 aa. [WO9608513-A1, 21-MAR-1996]	1..1602 1..1581	848/1620 (52%) 1121/1620 (68%)	0.0
AAM39043	Human polypeptide SEQ ID NO 2188 - <i>Homo sapiens</i> , 4618 aa. [WO200153312-A1, 26-JUL-2001]	627..2194 2901..4616	544/1741 (31%) 834/1741 (47%)	0.0
AAW18824	Human restrictin - <i>Homo sapiens</i> , 1358 aa. [US5635360-A, 03-JUN-1997]	484..1414 188..1107	338/935 (36%) 528/935 (56%)	0.0

In a BLAST search of public sequence databases, the NOV35a protein was found to have homology to the proteins shown in the BLASTP data in Table 35E.

Table 35E. Public BLASTP Results for NOV35a				
Protein Accession Number	Protein/Organism/Length	NOV35a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P24821	Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin) (GMEM) (JI) (Miotendinous antigen) (Glioma-associated-extracellular matrix antigen) (GP 150-225) (Tenascin-C) (TN-C) - <i>Homo sapiens</i> (Human), 2201 aa.	1..2199 1..2201	2194/2201 (99%) 2198/2201 (99%)	0.0
JQ1322	tenascin precursor - mouse, 2019 aa.	1..1796 1..1791	1282/1807 (70%) 1453/1807 (79%)	0.0
Q64706	Tenascin C precursor - <i>Mus musculus</i> (Mouse), 2019 aa.	1..1796 1..1791	1277/1807 (70%) 1449/1807 (79%)	0.0

Q29116	Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin) (GMEM) (JI) (Miotendinous antigen) (Glioma-associated- extracellular matrix antigen) (GP 150-225) (Tenascin-C) (TN-C) (P230) - <i>Sus scrofa</i> (Pig), 1746 aa.	1..1528 1..1521	1050/1532 (68%) 1213/1532 (78%)	0.0
P10039	Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin) (GMEM) (JI) (Miotendinous antigen) (Glioma-associated- extracellular matrix antigen) (GP 150-225) - <i>Gallus gallus</i> (Chicken), 1808 aa.	1..1602 1..1579	849/1618 (52%) 1123/1618 (68%)	0.0

Pfam analysis predicts that the NOV35a protein contains the domains shown in Table 35F.

Table 35F. Domain Analysis of NOV35a			
Pfam Domain	NOV35a Match Region	Identities/ Similarities for the Matched Region	Expect Value
EGF	185..216	10/48 (21%) 27/48 (56%)	0.34
EGF	251..278	13/47 (28%) 24/47 (51%)	0.51
EGF	283..309	12/47 (26%) 22/47 (47%)	0.0055
EGF	314..340	12/47 (26%) 21/47 (45%)	0.076
EGF	345..371	9/47 (19%) 20/47 (43%)	0.93
EGF	376..402	13/47 (28%) 22/47 (47%)	0.0026
EGF	407..433	14/47 (30%) 25/47 (53%)	0.0014
EGF	469..495	13/47 (28%) 22/47 (47%)	0.0049
EGF	500..526	13/47 (28%) 22/47 (47%)	0.0023
EGF	531..557	12/47 (26%) 23/47 (49%)	0.007

EGF	562..588	11/47 (23%) 24/47 (51%)	0.0033
EGF	593..619	12/47 (26%) 24/47 (51%)	0.023
fn3	622..700	29/85 (34%) 58/85 (68%)	5.5e-15
fn3	711..794	24/87 (28%) 65/87 (75%)	2.6e-13
fn3	802..881	26/85 (31%) 66/85 (78%)	1.9e-15
fn3	892..973	35/87 (40%) 65/87 (75%)	4.1e-19
fn3	984..1061	30/84 (36%) 65/84 (77%)	4.3e-16
fn3	1073..1156	26/87 (30%) 65/87 (75%)	2.8e-14
fn3	1164..1242	23/85 (27%) 58/85 (68%)	3.4e-13
fn3	1255..1334	26/85 (31%) 65/85 (76%)	3.4e-15
fn3	1346..1429	21/87 (24%) 64/87 (74%)	3.6e-13
fn3	1437..1513	20/85 (24%) 56/85 (66%)	8e-08
fn3	1528..1607	22/85 (26%) 58/85 (68%)	3.2e-11
fn3	1619..1698	21/85 (25%) 61/85 (72%)	2e-12
fn3	1709..1787	29/84 (35%) 58/84 (69%)	4.4e-17
fn3	1798..1875	23/84 (27%) 60/84 (71%)	8.5e-14
fn3	1886..1963	31/84 (37%) 60/84 (71%)	3.3e-19
fibrinogen_C	1979..2187	121/272 (44%) 208/272 (76%)	2.1e-134

Example 36.

The NOV36 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 36A.

Table 36A. NOV36 Sequence Analysis			
	SEQ ID NO: 187	4077 bp	
NOV36a, CG56054-01 DNA Sequence	GGAGCGGCGGGCGGGCGGGAGGGCTGGCGGGGCGAACGTCTGGGAGACGTCTGAAAGACC AACGAGACTTTGGAGACCAGAGACGCGCCTGGGGGGACCTGGGGCTTGGGGCGTGCAGAGA TTTCCCTTGCATTTCGCTGGGAGCTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCC GCGACCCTTGGGGGGCCTCCGGGATTGCTACCTTTTGGCTCCCTGCTCGTCAACTGCTG TCTTCTCAGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGCCAAGGAGGGCG AGCCAGGCAGCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCC AGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCAGCAGGCGAATC GCACTGGAGGCCTCTTCGCTTGCCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGG ACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACCAGTGGTTGGGAGTCA GTGTTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATATGAGGCAA GGCAGCGAGTGGACAGATCCTGGAGACGCGGGATATGATTGGTCGCTGCTTTGTGCTCA GCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGAC GCCCCAAGGCCATGAACAATTTGGGTTCTGCCAGCAGGGCACAGCTGCCGCCCTTCTCCC CTGATAGCCACTACCTCTCTTTGGGGCCCCAGGAACCTATAATTGGAAGGGGTTGCTTT TTGTGACCAACATGTAGACTCAGACCCCGACCGAGCTGGTGTATAAACTTTGGACCCTG CTGACCGGCTCCAGGACCAGCCGGAGACTTGGCCCTCAATAGTACTTAGGCTTCTCTA TTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAGAGCTGAGCTTTGTGGCTGGAGCCCCC GCGCCAACCACAAGGGTGTGTGTTATCCTGCGCAAGGACAGCGCCAGTGCCTGGTGC CCGAGGTTATGCTGTCTGGGAGCGCCTGACCTCCGGCTTTGCTACTCACTGGCTGTGG CTGACCTCAACAGTGTAGGCTGGCCAGACCTGATAGTGGGTGCCCCCTACTTCTTTGAGC GCCAAGAAGAGCTGGGGGGTGTGTGTATGTGTACTTGAACCAAGGGGGTCACTGGGCTG GGATCTCCCCTCTCCGGCTCTGCGGCTCCCCTGACTCCATGTTTCGGGATCAGCCTGGCTG TCCTGGGGGACCTCAACCAAGATGGCTTTCAGATATTGCAGTGGGTGCCCTTTGATG GTGATGGGAAAGTCTTCATCTACCATGGGAGCAGCCTGGGGGTTGTGCGCAAACTTCAC AGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGGCTACTCCCTGTGAGGCAGCT TGGATATGGATGGGAACCAATACCTGACCTGCTGGTGGGCTCCCTGGCTGACACCGCAG TGCTCTTCAGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAA GCATCGACCTGGAGCAGCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGG TCTGTTTCAGCTACATTGCAGTCCCCAGCAGCTATAGCCCTACTGTGGCCCTGGACTATG TGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGTTCCCGTGTGACGTTCTCTGA GCCGTAACTGGAAGAACCAAGCACCAGGCCTCGGGCACCCTGTGGCTGAAGCACCAGC ATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTC GGGCCATTGTAGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTC CTGGCCAGGGGCTGCCTCCAGTGGCCCCATCCTCAATGCCACCAGCCAGCACCAGC GGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGACAAGATCTGCCAGGCAATC TGCAGCTGGTCCACGCGCGTCTGTACCCGGGTGAGCGACACGGAATTCACACCTCTGC CCATGGATGTGGATGGAACAACAGCCCTGTTTGCACTGAGTGGGCAGCCAGTCATTGGCC TGGAGCTGATGGTCACCAACTGCCATCGGACCCAGCCAGCCAGCCAGGCTGATGGGGATG ATGCCCATGAAGCCAGCTCCTGGTCATGCTTCTGACTCACTGCACACTCAGGGGTCC GGGCCCTGGACCCTGCGGAGAAGCCACTCTGCCTGTCCAATGAGAATGCCTCCATGTG AGTGTGAGCTGGGGAACCCATGAAGAGAGGTGCCAGGTCACTTCTACCTCATCCTTA GCACCTCCGGGATCAGCATTGAGACCACGGAACGGAGGTAGAGCTGCTGTTGGCCACGA TCAGTGAGCAGGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTCATTGAGCTGCCAC TGTCCATTGCAGGAATGGCCATTCCCCAGCAACTCTTCTTCTCTGGTGTGGTGAGGGGCG AGAGAGCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGTCAAGTATGAGGTACGGTGT CCAACCAAGGCCAGTCGCTCAGAACCCTGGGCTCTGCCTTCTCAACATCATGTGGCCTC ATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGC AGGGGCTTGGGCAGAAAGGCTTTGCTCTCCAGGCCCAACATCCTCCACTGGATGTGG ACAGTAGGGATAGGAGGGCGCGGAGCTGGAGCCACTGAGCAGCAGGAGCCTGGTGAGC GGCAGGAGCCGACGATGCTGGTGGCCAGTGTCTCTGCTGAGAAGAGAAAACATCA CCCTGGACTGCGCCCGGGCACGGCCAACTGTGTGGTGTTCAGTGCCCACTCTACAGCT TTGACCGCGCGGCTGTGCTGCATGTCTGGGGCCGTCTCTGGAACAGCACCTTTCTGGAGG AGTACTCAGCTGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCCT CCATAAAGAACTTGATGCTCCGAGATGCCTCCACAGTGATCCAGTGATGGTATGCTTGG ACCCCAATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGTCACTCCTCGGCTGTACTGG CTGGGCTGTGGTGTAGCACTGCTGGTGTGCTCCTGTGGAAGATGGGATTCTTCAAAC GGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTCGGGAAG		

	ACCGACAGCAGTTCAAGGAGGAGAAGACGGGCACCATCTGAGGAACAACCTGGGGCAGCC CCCGCGGGAGGGCCCGGATGCACACCCCATCTGGCTGCTGACGGGCATCCCGAGCTGG GCCCCGATGGGCATCCAGGGCCAGGCACCGCCTAGGTTCCTCATGTCAGCCTGGCCTGT GGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGTGGGCT GCTGGTGTGCGATCAAGATTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCAC CCACAAGAAGCTCTCCCAACCACTTCCCTTAGAGTGCTGTGAGATGAGAGTGGGTAAA TCAGGGACAGGGCCATGGGGTAGGGTGAGAAGGGCAGGGGTGTCTGATGCAAGGTGGG GAGAAGGGATCCTAATCCCTTCTCTCCATTACCCCTGTGTAAAGAGACCCCAAGGACC TGCCTCCCCGGAAGTGCCTTAACCTAGAGGGTCGGGGAGGAGGTTGTGTCACTGACTCAG GCTGCTCCTTCTCTAGTTTCCCTCTCATCTGACCTTAGTTTGTGCTGCCATCAGTCTAGTG GTTTCGTGGTTTCGTCTATTATTAATAAAATATTGAGAACAAAAA		
	ORF Start: ATG at 162		ORF Stop: TAG at 3573
	SEQ ID NO: 188	1137 aa	MW at 124286.2kD
NOV36a, CG56054-01 Protein Sequence	MAGARSRDPWGASGICYLFGSLLEVLLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVIDIQGADMQKESKE NQWLGVSVRSQGPGGKIVTCAHRYEARQVRDQILETRDMIGRCFVLSQDLAIRDELDDGGE WKFCEGRPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGLLFVTNIDSSDPDQLV YKTLDPADRLPGPAGDLALNSYLGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKD SASRLVPEVMLSGERLTSGFGYSLAVADLNSDGPDLIVGAPYFFERQEEELGGA VVYVYN QGGHWAGISPLRLCGSPDSMFGISLAVLGDNLQDGFPIAVGAPFDGDGKVFYIYHSSLG VVAKPSQVLEGEAVGIKSFGYSLSGSLMDGNQYPDLLVGS LADTAVLFRARPILHVSHE VSIAPRSIDLEQPN CAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTRRLRGQV PRVTFLSRNLEEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYSLQTP RLRRQAPGQGLPPVAPILNAHQ PSTQRAEIHFLKQCGGEDKIQSNLQLVHARFCTRVSD TEFQPLPMDVDGTTALFALSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDS LHYSGVRALDPAEKPLCLSNENASHVECELGNPMKRG AQVTFYLILSTSGIS IETTELEV ELLLATISEQELHPVSARARVFIELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKV KYEVTVSNGQSLRTLGSFAFLNIMWPHEIANGKWL LYPMQVELEGGQGPQGKLCSPRPN ILHLDVDSRDRRRRELEPPEQEPGERQEPMSWVPVSSAEKKNITLDCARGTANC VVF SCPLYSFDRAAVLHVWGR LWNSTFLEEYSAVKSLEVI VRANITVKSSIKNMLRDASTVI PVMVYLDPM AVVAEGVPWWVILLAVLAGLLVLALLVLLWKMGFFKRAKHPEATVPQYHA VKIPREDRQQFKEEKTGTILRNWGS PRREGPD AHPILAADGHP ELGPDGHPGP GTA		
	SEQ ID NO: 189	2564 bp	
NOV36b, CG56054-03 DNA Sequence	GGAGCGCGGGCGGGCGGGAGGGCTGGCGGGGCGAACGCTGCGGAGACGCTGAAAGACC AACGAGACTTTGGAGACCAGAGACGCGCCTGGGGGGACCTGGGGCTTGGGGCGTGCAGAG TTTCCCTTGCATTGCTGGGAGACTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCC GCGACCTTGGGGGGCCTCCGGGATTTGCTACCTTTTGGCTCCCTGCTCGTGAAGTGC TCTTCTACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGCGCAAGGAGGGCG AGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCTGCACCGGCAGTTGCAGCCCCGACCC AGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCAGCAGGCGAATC GCACTGGAGGCCCTCTCGCTTGCCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGG ACATCGACCAGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACCAAGTGGTGGGAGTCA GTGTTGCGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATATGAGGCCAA GGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTGCCTGCTTTGTGCTCA GCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGAATGGAAGTTCTGTGAGGGAC GCCCCAAGGCCATGAACAATTTGGGTTCTGCCAGCAGGGCACAGCTGCCGCTTCTCCC CTGATAGCCACTACCTCCTTTTGGGGCCCCAGGAACCTATAATTGGAAGGGGTTGCTTT TTGTGACCAACATTGATAGCTCAGACCCCGACCAGCTGGTGTATAAACTTTGGACCCTG CTGACCGGCTCCCAGGACCAGCCGAGACTTGGCCCTCAATAGCTACTTAGGCTTCTCTA TTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAGAGCTGAGCTTTGTGGCTGGAGCCCCC GCGCCAACCACAAGGGTGCTGTGTTATCCTGCGCAAGGACAGCGCCAGTCGCTGGTGC CCGAGGTTATGCTGTCTGGGAGCGCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGG CTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGGTGCCCTACTTCTTTGAGC GCCAAGAAGAGCTGGGGGGTGCTGTGTATGTGTACTTGAACCAGGGGGGTCACTGGGCTG GGATCTCCCTCTCCGGCTCTGCAACTCCCCGCACTCCATGTTCCGGATCAGCCTGGCTG TCCTGGGGGACCTCAACCAAGATGGCTTCCAGATATTGCAGTGGGTGCCCCCTTTGATG GTGATGGGAAAGTCTTCACTACCATGGGAGCAGCTGGGGGTTGTGCCAAACCTTCAC AGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGGCTACTCCCTGTCAGGCAGCT		

	TGGATATGGATGGGAACCAATACCTTGACCTGCTGGTGGGCTCCCTGGCTGACACCGCAG TGCTCTTCAGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAA GCATCGACCTGGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGG TCTGTTTCAGCTACATTGCAGTCCCCAGCAGCTATAGCCCTACTGTGGCCCTGGACTATG TGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGTTCCCGTGTGACGTTCTCTGA GCCGTAACCTGGAAGAACCACCAAGCACCAGGCCTCGGGCACCCTGTGGCTGAAGCACCAGC ATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTC GGGCCATTGTAGTGACCTTGTCTACAGTCTCCAGACCCCTCGGTCCGGCGGGAGGGCC CGGATGCACACCCCATCTGGCTGCTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATC CAGGGCCAGGCACCGCCTAGGTTCCCATGTCCAGCCTGGCCTGTGGCTGCCCTCCATCC CTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGTGGGCTGCTGGTGTGCGATCA AGATTGCGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAACTCCTC CCACCCAACTTCCCCTTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCA TGGGGTAGGGTGAAGAGGGCAGGGGTGTCTGATGCAAAGGTGGGGAGAAGGGATCCTAA TCCCTTCTCTCTCCATTACCTGTGTAAACAGGACCCCAAGGACCTGCCTCCCGGAAGT GCCTTAACCTAGAGGTCTGGGGAGGAGGTGTGTCTACTGACTCAGGCTGCTCTTCTCTA GTTTCCTCTCTCATCTGACCTTAGTTTGTGCTGCCATCAGTCTAGTGGTTTCGTGGTTTCGT CTATTTATTAATAAATATTTGAGAACAATAAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 162		ORF Stop: TAG at 2058
	SEQ ID NO: 190	632 aa	MW at 68332.4kD
NOV36b, CG56054-03 Protein Sequence	MAGARSRDPWASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVIDIQADMQKESKE NQWLGVSVRSQPGGKIVTCAHRYEARQVRVDQILETRDMIGRCFVLSQDLAIRDELDDGGE WKFCGRPQGHEQFGFCQQTAAAFSPDSHYLLFGAPGTYNWKGLLFVTNIDSSDPDQLV YKTLDPADRLPGPADLALNSYLGFSDSGKGLVRAEELSFAVAGAPRANHKGAVVILRKD SASRLVPEVMSLGERLTSGFGYSLAVADLNSDGWPDLLVGPYFFERQEEELGGAVVYVYN QGGHWAGISPLRLCNSPHSMFGISLAVLGDNLQDGFDPDIAVGAPFDGDKVFIYHGSSLG VVAKPSQVLEGEAVGKISFGYSLSGSLMDGNQYPDLLVGSADTAFLFRARPILHVSHE VSIAPRSIDLEQPNACGHSVCVDLRVCFYSYIAVPSSYSPTVALDYVLDADTDRLRLRGV PRVTFLSRNLEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYSLQTP RLRREGPDAPILAADGHPGLGPDGHPGPGTA		
	SEQ ID NO: 191	2017 bp	
NOV36c, CG56054-04 DNA Sequence	GGAGCGGCGGGCGGGCGGGAGGGCTGGCGGGCGAACCTCTGGGAGACGTCTGAAAGACC AACGAGACTTTGGAGACCAGAGACGCGCCTGGGGGACCTGGGGCTTGGGGCTGCGAGA TTTCCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCC GCGACCCTTGGGGGCTCCTCGGGATTGCTACCTTTTGGCTCCTGCTCGTGAACCTGC TCTTCTCAGGGGTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGGCGAAGGAGGGCG AGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGACGCCCCGACCCC AGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCAGACGGCAGATC GCACTGGAGGCCTCTTCGCTTGCCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGG ACATCGACCAGGAGCTGATATGCAAAGGAAAGCAAGGAGAACAGTGGTTGGGAGTCA GTGTTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATATGAGGCAA GGCAGCGAGTGGACAGATCCTGGAGACGCGGGATATGATTGGTCTGCTTTGTGCTCA GCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGAC GCCCCAAGGCCATGAACAATTTGGGTTCTGCCAGCAGGGCAGAGCTGCCGCTTCTCCC CTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAACCTATAATTGGAAGGGGTGCTTT TTGTGACCAACATTGATAGCTCAGACCCGACCAGCTGGTGTATAAACTTTGGACCTG CTGACCGGCTCCAGGACCAGCCGAGACTTGGCCCTCAATAGCTACTTAGGCTTCTCTA TTGACTCGGGGAAGGTCTGCTGCTGCAGAAAGAGCTGAGCTTGTGGCTGGAGCCCCC GCGCAACCACAAGGTGCTGTGGTATCCTGCGCAAGGACAGCGCCAGTCGCTGGTGC CCGAGGTTATGCTGTCTGGGGAGCGCCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGG CTGACCTCAACAGTGTGCTGGCCAGACCTGATAGTGGGTGCCCCCTACTTCTTTGAGC GCCAAGAAGAGCTGGGGGGTGTGTGTATGTGTACTTGAACAGGGGGGTCACTGGGCTG GGATCTCCCCCTCTCCGGCTCTCGGCTCCCTGACTCCATGTTCCGGATCAGCTGGCTG TCTGGGGGACCTCAACCAAGATGGCTTCCAGATATTGCAAGTGGGTGCCCCCTTTGATG GTGATGGGAAGTCTTCTATCTACCATGGGAGCAGCCTGGGGTGTGCGCAAGCCTTAC AGGTGCTGGAGGGCAGGCTGTGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCC AGGCACCGCCTAGGTTCCCATGTCCAGCCTGGCTGTGGCTGCCCTCCATCCCTTCCCC		

	AGAGATGGCTCCTTGGGATGAAGAGGGTAGAGTGGGCTGCTGGTGTGCGCATCAAGATTG GCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAACTCTCCACCCA ACTTCCCCTTAGAGTGCTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTA GGGTGAGAAGGGCAGGGGTGTCTGTATGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTC CTCTCCATTACCCCTGTGTAACAGGACCCCAAGGACCTGCCCTCCCGAAGTGCCTTAA CCTAGAGGGTCGGGGAGGAGTTGTGTCACTGACTCAGGCTGCTCCTTCTAGTTCCCT CTCTCATCTGACCTTAGTTTGCTGCCATCAGTCTAGTGGTTTCGTGGTTTCGTCTATTTA TTAAAAATATTGAGAACAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 162		ORF Stop: TGA at 1764
	SEQ ID NO: 192	534 aa	MW at 57440.7kD
NOV36c, CG56054-04 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVLELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVIDIDQADMOKESKE NQWLGVSVRSGPGGKIVTCAHRYEARQVRDQILETRDMIGRCFVLSQDLAIRDELDDGGE WKFCEGRPQGHQFGFCQOGTAAAFSPDSHYLLFGAPGTYNWKGLLFVTNIDSSDPDQLV YKTLDPADRLPGFAGDLALNSYLGFSDSGKGLVRABELSFVAGAPRANHKGAVVILRKD SASRLVPEVMLSGERLTSFGYSLAVADLNSDGWPDLI VGAPYFFERQELGGAVVYVLYN QGGHWAGISPLRLCGSPDSMFGISLAVLGDNLQDGLPDI AVGAPFDGDGVFIYHGSLSL VVAKPSQVLEGEAVGIPSWAPMGIQQAAPPFP CPSLACGCPPSLPQRWLLGMKRVEWAA GVASRFGRIGFLRGTDLSHPQELLPNPFLECCEMRVGKSGTGPWGRVRRAGVS		
	SEQ ID NO: 193	999 bp	
NOV36d, CG56054-05 DNA Sequence	ATGGCCGGGGCTCGGAGCCGCGACCTTGGGGGGCTCCGGGATTGTGACCTTTTGGC TCCCTGCTCGTCGAAGTGTCTTCTCAGGGGCTGTGCGCTTCAATCTGGACGTGATGGGT GCCTTGCGCAAGGAGGGCGAGCCAGGCAGCTCTTCGGCTTCTGTGGCCCTGCACCGG CAGTTGCAGCCCCGACCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTT CCTGGGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTGGCCGTTGAGCCTGGAGGAG ACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAGGAG AACCAGTGGTTGGGAGTCACTGTTCCGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGT GCACACCCATCCTGGCTGTGACGGGCATCCGAGCTGGGCCCGCATGGGCATCCAGGG CCAGGCACCGCCTAGGTTCCCATGTCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCC CCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGTGGGCTGCTGGTGTGCGCATCAAGATT TGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAACTCTCCACCC CAACTTCCCCTTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGG TAGGGTGAGAAGGGCAGGGGTGTCTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCT TCCTTCCCATTACCCCTGTGTAACAGGACCCCAAGGACCTGCCCTCCCGGAAGTGCCCT AACCTAGAGGGTCGGGGAGGAGGTTGTGTCACTGACTCAGGCTGCTCCTTCTTAGTTTC CCCTCTCATCTGACCTTAGTTTGCTGCCATCAGTCTAGTGGTTTCGTGGTTTCGTCTATT TATTAATAATATTGAGAACAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 1		ORF Stop: TAG at 493
	SEQ ID NO: 194	164 aa	MW at 17332.5kD
NOV36d, CG56054-05 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVLELLFSRAVAFNLDVMGALRKEGEPGSLFGF SVAL HRQLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVIDIDQGA DMQK ESKENQWLGVSVRSGPGGKIVTCAHPILAADGHPGLGPDGHPGPGTA		
	SEQ ID NO: 195	2701 bp	
NOV36e, CG56054-06 DNA Sequence	GGAGCGGGCGGGCGGGCGGGAGGGCTGGCGGGGCGAACGTCTGGGAGACGTCTGAAAGACC AACGAGACTTTGGAGACCAGAGACGCGCCTGGGGGGACCTGGGGCTTGGGGCGTGCGAGA TTTCCCTTGCAATTCGCTGGGAGCTCGCGCAGGGATCGTCCATGGCCGGGGCTCGAGCC GCGACCCCTGGGGGGCCTCCGGGATTGCTACCTTTTGGCTCCCTGCTCGTGAAGTGC TCTTCTCAGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCTTGGCAAGGAGGGCG AGCCAGGCAGCCTCTTCGGCTTCTGTGGCCCTGCACCGGCAGTTGAGCCCTGGACTA TGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGTTCCCCGTGTGACGTTCT GAGCCGTAACTTGAAGAACCAAGCACCAGGCCTCGGGCACCCTGTGGCTGAAGCACA GCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCT TCGGGCCATTGTAGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGC TCCTGGCCAGGGGCTGCCTCAGTGGCCCCATCCTCAATGCCACCCAGCCAGCACCCA		

	<p>CGCGGCGAGAGATCCACTTCTCTGAAGCAAGGCTGTGGTGAAGACAAGATCTGCCAGAGCAA TCTGCAGCTGGTCCACGCCCGCTTCTGTATCCCGGGTCAGCGACACGGAAATTCACCACTCT GCCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGAAGTGGGAGCCAGTCATTGG CCTGGAGCTGATGGTCACCAACCTGCCATCGGACCCAGCCAGCCAGCCAGGCTGATGGGGA TGATGCCCATGAAGCCAGCTCCTGGTCATGCTTCTGACTCACTGACACTACTCAGGGGT CCGGGCCCTGGACCCTGCGGAGAAGCCACTTGCCTGTCCAATGAGAATGCCTCCCATGT TGAGTGTGAGCTGGGGAACCCATGAAGAGAGGTGCCAGGTACCTTCTACCATCATGTGGCC TAGCACCTCCGGGATCAGCATTGAGACCACGGAAGTGGAGGTAGAGCTGCTGTTGGCCAC GATCAGTGAGCAGGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTATTGAGCTGCC ACTGTCCATTGCAGGAATGCCATTCCCCAGCAACTCTTCTCTCTGGTGTGGTGAGGGG CGAGAGGACCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGTCAAGTATGAGGTACGGT TTCCAACCAAGGCCAGTCGCTCAGAAACCTGGGCTCTGCCTTCTCAACATCATGTGGCC TCATGAGATTGCCAATGGGAAGTGGTGTCTGTACCCAATGCAGGTTGAGCTGGAGGGCGG GCAGGGGCTTGGGCAGAAAGGGCTTTGCTCTCCAGGCCCAACATCCTCCACCTGGATGT GGACAGTAGGGATAGGAGGCGGCGGAGCTGGAGCCACCTGAGCAGCAGGAGCCCTGGTGA CGCGCAGGAGCCAGCATGTCTGGTGGCCAGTGTCTCTGCTGAGAAGAAGAAAAAATC CACCTTGGACTGCGCCCGGGGACGGCCAAGTGTGGTGTTCAGTGTGCCACTTACAG CTTTGACCGCGCGGCTGTGTGCATGCTGTGGGCGCTCTCTGGAACAGCACCTTTCTGGA GGAGTACTCAGCTGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCAGTGAAAGTC CTCCATAAAGAAGTGTGATGCTCCGAGATGCCTCCACAGTGATCCAGTGATGGTATACTT GGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGTCACTCCTCTGGCTGTACT GGCTGGGTGCTGGTGCTAGCACTGCTGGTGCTGCTCCTGTGGAAGATGGGATCTTCAA ACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTCGGGA AGACCGACAGCAGTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAAGTGGGGCAG CCCCCGCGGGAGGGGCCCGGATGCACACCCCATCCTGGTCTGACGGGCATCCGAGCT GGGCCCCGATGGGCATCCAGGGCCAGGCACCGCCTAGGTTCCCATGTCCAGCCTGGCT GTGGCTGCCCTCAGCTCCCTTCCCGAGAGATGGTCTCTGGGATGAAGAGGGTAGAGTGG CTGCTGGTGTCCATCAAGATTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCC ACCCACAAGAACTCCTCCACCCAAGTCTCCCTTAGAGTGCTGTGAGATGAGAGTGGGTA AATCAGGGACAGGGCCATGGGGTAGGGTGAGAAGGGCAGGGGTGTCTGATGCAAGAGTG GGGAGAAGGGATCCTAATCCCTTCTCTCCCATTCACCTGTGTAAACAGGACCCCAAGGA CCTGCTCCCGGAAGTGCCTTAACCTAGAGGGTCCGGGAGAGGTTGTGTCACTGACTC AGGCTGCTCCTTCTCTAGTTTCCCTCTCATCTGACCTTAGTGTGCTGCCATCAGTCTAG TGGTTTCGTGGTTTCGTCTATTTATTAATAAATATTGAGAACAAAAAAAAAAAAAAAAA A</p>		
	ORF Start: ATG at 162		ORF Stop: TAG at 366
	SEQ ID NO: 196	68 aa	MW at 7433.6kD
NOV36e, CG56054-06 Protein Sequence	MAGARSRDPWAGSGICYLFGSLLEVLLFSRAVAFNLDVMGALRKEGEPGSLFGPSVALHR QLQPWTMC		
	SEQ ID NO: 197	1131 bp	
NOV36f, CG56054-07 DNA Sequence	GGAGCGGCGGGCGGGCGGGAGGGCTGGCGGGGCGAACGCTGGGAGACGCTGAAAGACC AACGAGACTTTGGAGACCAGAGACGCGCCTGGGGGGACCTGGGGCTTGGGGCGTGCGAGA TTTCCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCC CGACACCTTGGGGGGCTCGGGATTGCTACCTTTTGGCTCCCTGCTCGTCAAGTGC TCTTCTACGGGCTGTGCCTTCAATCTGACGCTGATGGGTGCTTGCAGGAAGGGCG AGCCAGGCAGCCTCTTCGGCTTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCC AGAGCTGGCTGCTGGTGGTGCTCCCCAGGCCCTGGCTTCTCTGGGCAGCAGGCGAATC GCACTGGAGGCCTCCGTGCCCCAGTACCATGCGGTGAAGATTCTCTGGGAAGACCGACAG CAGTTCAAGGAGGAGAAGACGGGCACCATCTGAGGAACAAGTGGGGCAGCCCCGGCGG GAGGGCCCGGATGCACACCCATCTGGCTGCTGACGGGCATCCGAGCTGGGCCCGCAT GGGCATCCAGGGCCAGGCACCGCCTAGGTTCCCATGTCCAGCCTGGCCTGTGGCTGCC TCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGTGGGCTGCTGGTGT CGCATCAAGATTGGGAGGATCGGCTTCTCAGGGGCACAGACCTTCCCACCCACAAGA ACTCCTCCCAACCACTTCCCTTAGAGTGCTGTGAGATGAGAGTGGGTAATCAGGGAC AGGGCCAGTGGGTAGGTTGAGAAGGGCAGGGGTGTCTGATGCAAGGTGGGAGAGGG ATCCTAATCCCTTCTCTCCCATTCACCTGTGTAAACAGGACCCCAAGGACCTGCCTCCC		

	CGGAAGTGCCTTAACCTAGAGGGTCGGGGAGGAGGTGTGTCACTGACTCAGGCTGCTCC TTCTCTAGTTTCCCTCTCATCTGACCTTAGTTTGCTGCCATCAGTCTAGTGGTTTCGTG GTTTCGTCTATTTATTAATAAATATTTGAGAACAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 162		ORF Stop: TGA at 573
	SEQ ID NO: 198	137 aa	MW at 14203.9kD
NOV36f, CG56054-07 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLRAPVPCGEDSSGRPTAVQGGEDGHHPEEQ LGQPPAGGPGCTPHPGC		
	SEQ ID NO: 199	2175 bp	
NOV36g, CG56054-08 DNA Sequence	GGAGCGGCGGGCGGGCGGGAGGGCTGGCGGGCGAACGTCTGGGAGACGTCTGAAAGACC AACGAGACTTTGGAGACCAGAGACGCGCCTGGGGGGACCTGGGGCTTGGGGCGTGCAGAG TTCCCTTGCAATTCGCTGGGAGCTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCC GCGACCCCTGGGGGGCTCCGGGATTTGCTACCTTTTGGCTCCCTGCTCGTCAACTGC TCTTCTCACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGCAGGAGGGCG AGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCC AGAGCTGGCTGCTGGTGGGTGCTCCCGAGGCCCTGGCTCTTCTGGGCAGCAGGCGAATC GCACTGGAGGCCTCTTCGCTTGCCCGTTGAGCCTGGAGGAGACTGACTGTACAGAGTGG ACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACCAGTGGTTGGGAGTCA GTGTTCCGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATATGAGGCAA GGCAGCGAGTGGACAGATCCTGGAGACGCGGATATGATTGGTTCGCTGCTTTGTGCTCA GCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGAATGGAAGTTCTGTGAGGGAC GCCCCAAGGCCATGAACAATTGGGTTCTGCCAGCAGGGCACAGCTGCCGCCCTTCTCCC CTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAACCTATAATTGGAAGGGGTGCTTT TTGTGACCAACATTGATAGCTCAGACCCGACAGCTGGTGTATAAACTTTGGACCTG CTGACCGGCTCCCAGGACACGCGGAGACTTGGCCCTCAATAGCTACTTAGGCTTCTCTA TTGACTCGGGGAAAGGTCTGGTGCCTGCAGAGAGCTGAGCTTTGTGGCTGGAGCCCCC GCGCCAACCACAAGGGTGTGTGGTTCATCTGCGCAAGGACAGCGCCAGTGCCTGGTGC CCGAGGTTATGCTGTCTGGGGAGCGCCTGACCTCCGGCTTTGGCTACTACTGGCTGTGG CTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGGTGCCCTACTTCTTTGAGC GCCAAGAAGAGCTGGGGGGTGTGTGTATGTGTAAGTGAACAGGGGGCTACTGGGCTG GGATCTCCCTCTCCGGCTCTGCGGCTCCCTGACTCCATGTTCCGGATCAGCCTGGCTG TCCTGGGGGACCTCAACCAAGATGGCTGTGGTGGCAGAGGAGTGCCCTGGTGGTTCATC CTCCTGGCTGTACTGGCTGGGCTGCTGGTGTAGCACTGCTGGTGTGCTCCTGTGGAAG ATGGGATTCTTCAAACGGGCGAAGCACCCGAGGCCACCGTGCCCCAGTACCATGCGGTG AAGATTCTCGGGAAGACCGACAGCAGTTCAAGGAGGAGAAGACGGGCACCATCTGAGG AACAACCTGGGCGAGCCCCGCGGGAGGGCCCCGATGCACACCCCATCTGGCTGTGAC GGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCTAGGTTCCCATG TCCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCAGAGATGGCTCCTTGGGATGAA GAGGGTAGAGTGGGCTGCTGGTGTGCGATCAAGATTGGCAGGATCGGCTTCTCAGGGG CACAGACCTCTCCACCCACAAGAACTCTCCACCCAACTTCCCTTAGAGTGCTGTGA GATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGGTGAGAAGGGCAGGGGTGTC CTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTACCCCTGTGTAA CAGGACCCCAAGGACCTGCCTCCCGGAAGTGCTTAACCTAGAGGGTCGGGGAGGAGGT TGTGTCACTGACTCAGGCTGCTCCTTCTAGTTTCCCTCTCATCTGACCTAGTTTGC TGCCATCAGTCTAGTGGTTTCGTGGTTTCGTCTATTTATTAATAAATATTTGAGAACAAA AAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 162		ORF Stop: TGA at 1617
	SEQ ID NO: 200	485 aa	MW at 51430.2kD
NOV36g, CG56054-08 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVDIDQADMOKESKE NQWLGVSVRSQPGGKIVTCAHRYEARQVRDQILETRDMIGRCFVLSQDLAIRDELGGGE WKFCEGRPQGHEQFGFCQQTAAAFSPDSHYLLFGAPGTYNWGLLFVTNIDSSDPDQLV YKTLDPADRLPGPAGDLALNSYLGFSDSGKGLVRAEELSFVAGAPRANHKGAVVILRKD SASRLVPEVMSGERLTSFGYS LAVADLNSDGWPDILVGPYFFERQEELGGAVVYVLYN QGGHWAGISPLRLCGSPDSMFGISLAVLGDNLNQDCGGRRSALVGHPPGCTGWAAGASTA		

	GAAPVEDGILQTGEAPRGHRAPVPCGEDSSGRPTAVQGGEDGHHPEEQLGQPPAGGPGCT PHPGC		
	SEQ ID NO: 201	1458 bp	
NOV36h, CG56054-09 DNA Sequence	TTGGGGCGTGCGAGATTCCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGG CCGGGGCTCGGAGCCGCGACCTTGGGGGGCTCCGGGATTGTACTCTTTTGGCTCCC TGCTCGTCGAACTGCTCTTCTACGGGCTGTCGCCTTCAATCTGGACGTGATGGGTGCCT TGC GCAAGGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTGTGTGCCCTGCACCGGCAGT TGCAGCCCCGACCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTCTG GGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTGCCCGTTGAGCCTGGAGGAGACTG ACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACC AGTGGTTGGGAGTCAGTGTTCCGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCAC ACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTCT GCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGA AGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATTGGGTTCTGCCAGCAGGGCACAG CTGCCGCTTCTCCCTGATAGCCACTACCTCCTCTTGGGGCCCCAGGAACCTATAATT GGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGG ACGACGGTCCCTACGAGGCGGGGGAGAGAAGGAGCAGGACCCCGCCTCATCCCGGTCC CTGCGAACAGCACCTTTCTGGAGGAGTACTCAGCTGTGAAGTCCCTGGAAGTGATTGTCC GGGCCAACATCACAGTGAAGTCTCTCATAAAGAAGTGGATGCTCCGAGATGCCTCCACAG TGATCCCACTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGT GGGTCATCCTCCTGGCTGTACTGGCTGGGCTGCTGGTGTAGCACTGCTGGTGTGCTCC TGTGGAAGATGGGATTCTTCAAACGGGCGAAGCACCCCGAGGCCACCGTGCCTCCAGTACC ATGCGGTGAAGATTCTCGGGAAGACCGACAGCAGTTCAAGGAGGAGAAGACGGGACCA TCCTGAGGAACAACCTGGGGCAGCCCCCGCGGGAGGGCCCGGATGCACACCCCATCTGG CTGCTGACGGGCATCCCGAGCTGGGCCCGATGGGCATCCAGGGCCAGGCACCGCCTAGG TTCCCATGTCCAGCCTGGCTGTGGCTGCCCTCCATCCCTTCCCAGAGATGGCTCCTT GGGATGAAGAGGGTAGAGTGGGCTGCTGGTGTGCGATCAAGATTGGCAGGATCGGCTTC CTCATGGGCACAGACCTC		
	ORF Start: ATG at 57		ORF Stop: TAG at 1317
	SEQ ID NO: 202	420 aa	MW at 45990.1kD
NOV36h, CG56054-09 Protein Sequence	MAGARSRDPWAGSGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVIDIDQADMQKESKE NQWLGVSVRSQPGGKIVTCAHRYEARQVRVDQILETRDMIGRCFVLSQDLAIRDELDDGE WKFCGRPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGRTARVELCAQGSADLAH LDDGPYEAGGEKEQDRLIPVPANSTFLEEYSAVKSLEIVIRANITVKSSIKNMLRDAS TVIPVMVYLDPMVAEVPWVWILLAVLAGLLVLLVLLWKMFGFKRAKHPEATVPQ YHAVKIPREDRQQFKKEKTGTILRNWGSPPREGPAHPILAADGHPGLGDPGHPGTA		
	SEQ ID NO: 203	3595 bp	
NOV36i, CG56054-10 DNA Sequence	TTGGGGCGTGCGAGATTCCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGG CCGGGGCTCGGAGCCGCGACCTTGGGGGGCTCCGGGATTGTACTCTTTTGGCTCCC TGCTCGTCGAACTGCTCTTCTACGGGCTGTCGCCTTCAATCTGGACGTGATGGGTGCCT TGC GCAAGGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTGTGTGCCCTGCACCCGCACT TGCAGCCCCGACCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTCTG GGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTGCCCGTTGAGCCTGGAGGAGACTG ACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACC AGTGGTTGGGAGTCAGTGTTCCGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCAC ACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTC GCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGA AGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATTGGGTTCTGCCAGCAGGGCACAG CTGCCGCTTCTCCCTGATAGCCACTACCTCCTCTTGGGGCCCCAGGAACCTATAATT GGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGG ACGACGGTCCCTACGAGGCGGGGGAGAGAAGGAGCAGGACCCCGCCTCATCCCGGTCC CTGCCAACAGCTACTTTGGCTTCTCTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAG AGCTGAGCTTTGTGGCTGGAGCCCCCGGCCAACCAAGGGTGTGTGGTGCATCCTGC GCAAGGACAGCGCCAGTGCCTGGTGGCCGAGGTTATGCTGTCTGGGAGCGCCTGACCT CCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGGCCAGACCTGA TAGTGGGTGCCCCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGTGCTGTGTATGTGT		

	<p>ACTTGAACCAGGGGGTCACTGGGCTGGGATCTCCCTCTCCGGCTCTGCGGCTCCCTG ACTCCATGTTCCGGGATCAGCCTGGCTGTCTGGGGGACCTCAACCAAGATGGCTTTCCAG ATATTGCAGTGGGTGCCCCCTTTGATGGTGATGGGAAAGTCTTCATCTACCATGGGAGCA GCCTGGGGGTTGTGCGCAAACCTTCACAGGTGTGAGGGGCGAGGCTGTGGGCATCAAGA GCTTCGGCTACTCCCTGTGAGGAGCTTGGATATGGATGGGAACCAATACCTTGACCTGC TGGTGGGCTCCCTGGCTGACACCGCAGTGCTCTTCAGGGCCAGACCCATCCTCCATGTCT CCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCTGGAGCAGCCCACTGTGCTGGCG GCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCCCAGCAGCT ATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGG GCCAGGTTCCCCGTGTGACGTTCTGAGCCGTAACCTGGAAGAACCAAGCACCAGGCCT CGGCACCGTGTGGCTGAAGCACCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGC TCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTGTAGTGACCTTGCTTACAGTCTCC AGACCCCTCGGCTCCGGGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGCCCCATCC TCAATGCCCCACAGCCAGCAGCCAGCGGGCAGAGATCCACTTCCTGAAGCAAGGCTGTG GTGAAGACAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGG TCAGCGACACGGAATTCCAACCTCTGCCCATGGATGTGGATGGAACAACAGCCCTGTTTG CACTGAGTGGGCAGCCAGTCATTGGCTGGAGCTGATGGTCACCAACCTGCCATCGGACC CAGCCAGCCCCAGGCTGATGGGGATGATGCCCATGAAGCCAGCTCCTGGTCATGCTTC CTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCTGCGGAGAGCCACTCTGCC TGTCGAATGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCATGAAGAGAGGTG CCCAGGTCACCTTCTACCTCATCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAAC TGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCAGGAGCTGCATCCAGTCTCTGCAC GAGCCCGTGTCTTCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCCCCAGCAAC TCTTCTTCTCTGGTGTGGTGAGGGGCGAGAGAGCCATGCAGTCTGAGCGGGATGTGGGCA GCAAGGTCAAGTATGAGGTACCGTTTCCAACCAAGGCCAGTCGCTCAGAACCCTGGGCT CTGCCTTCTCAACATCATGTGGCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACC CAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCTGGGCAGAAAGGGCTTTGCTCTCCCA GGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGGGAGCTGGAGC CACCTGAGCAGCAGGAGCCTGGTGAGCGGCAGGAGCCAGCATGTCTGGTGGCCAGTGT CCTCTGCTGAGAAGAAGAAAAACATCACCTGGACTGCGCCCGGGGACGGCCAACTGTG TGGTGTTCAGCTGCCACTCTACAGCTTTGACCGCGGGCTGTGCTGCATGTCTGGGGCC GTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGCTGTGAAGTCCCTGGAAGTGATTG TCCGGGCCAACATCACAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGATGCCTCCA CAGTGATCCCAGTGATGTTACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCT GGTGGGTCACTCTCTGGCTGTACTGGCTGGGCTGTGGTGTAGCACTGCTGGTGTCTGC TCCTGTGGAAGTGTGGCTTCTTCCATCGGAGCAGCCAGAGCTCATCTTTCCACCAACT ATACCGGGCTGTCTGGCTGTGCAGCCTTCAGCCATGGAAGTGGGGGCTCAGGGCAGT TGGGATGGGATTCTTCAAACGGGCGAAGCACCCCGAGGCCACCGTGGCCAGTACCATGC GGTGAAGATTCTTCGGGAAGACCGACAGCAGTCAAGGAGGAGAAGACGGGCACCATCCT GAGGAACAATGGGGCAGCCCCGCGGGAGGGCCCGGATGCACACCCCATCTGGCTGC TGACGGGCATCCCGAGCTGGGCCCGGATGGGCATCCAGGGCCAGGCACCGCCTAG</p>
	<p>ORF Start: ATG at 57</p>
	<p>ORF Stop: TGA at 3423</p>
	<p>SEQ ID NO: 204</p>
	<p>1122 aa</p>
	<p>MW at 122352.9kD</p>
<p>NOV36i, CG56054-10 Protein Sequence</p>	<p>MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVIDQDADMQKESKE NQWLGVSVRSQPGGKIIVTCAHRYEARQVRDQILETRDMIGRCFVLSQDLAIRDELDDGE WKFCBGRPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGARVELCAQGSADLAH LDDGPYEAGGEKEQDPRILIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVI LRKDSASRLVPEVMSGERLTSGFGYSLAVADLNSDGWPDILVGPYFFERQEEELGGAVY VYLNQGGHWAGISPLRLCGSPDSMFGISLAVLGDNLQDGFDPDIAVGAPFDGDKGVFIYHG SSLGVVAKPSQVLEGEAVGKISFGYSLSGSLMDGNQYPDLLVGLSLADTAVLFRARPILH VSHEVSIAPRSIDLEQPNACGHSVCVDLRVCFYSYIAVPSSYSPTVALDYVLADTDRLR RGQVPRVTFLSRNLEEPKHQASGTVVLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYS LQTPRLRRQAPGQGLPPVAPILNAHQPSQRAEIHFLKQCGEDKICQSNLQLVHARFCT RVSDTEFQPLPMDVDGTTALFALSQGPVIGLELMVTNLPSPDPAQPQADGDDAHEAQLLV LPDSLHYSGVRALDPAEKPLCLSNENASHVECELGNPMKRGAVTFYILISTSGISIIET ELEVELLLATISEQELHPVSARARVFIELPLSIAGMAIPQQLFFSGVVRGERAMQSERDV GSKVKYEVTVSNNQGSRLTGSALFNIMWPHEIANGKWLLYPMQVELEGGQGPQKGLCS</p>

	PRPNILHLDVDSRDRRRRELEPPEQQEPGERQEPSMSWVPVSSAEKKKNITLDCARGTAN CVVFSCPLYSFDRAAVLHVWGRLLWNSTFLEEYSAVKSLEIVRANITVKSSIKNLMRLDA STVIPVMVYLDPMVAEVPWWVILLAVLAGLLVLALLVLLWKCGFFHRSSQSSSFPT NYHRACLAVQPSAMEVGGPGTVGWSSNGRSTPRPPCPSTMR		
	SEQ ID NO: 205	1034 bp	
NOV36j, CG56054-11 DNA Sequence	GGAGCGGCGGGCGGGCGGGAGGGCTGGCGGGGCGAACGTCTGGGAGACGTCTGAAAGACC AACGAGACTTTGGAGACCAGAGACGCGCCTGGGGGGACCTGGGGCTTGGGGCGTGCGAGA TTTCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCC GCGACCCTTGGGGGGCCTCCGGGATTGCTACCTTTTGGCTCCCTGCTCGTGAAGTGC TCTTCTACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGCAGCAAGGAGGGCG AGCCAGGCAGCCTCTTCGGCTTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCC AGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGACGACAGGCGAATC GCACTGGAGGCCTCTTCGTTGCCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGG ACATCGACCAGGGAGCTGATATGCAAAGGAAAGCAAGGAGAACCAGTGGTTGGGAGTCA GTGTCTCTGCTGAGAAGAAGAAAACATCACCTGGACTGCGCCCGGGGCACGGCCAAC TGTGTGGTGTTCAGCTGCCACTCTACAGCTTTGACCGCGCGGCTGTGTGTCATGTCTGG GGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGCTGTGAAGTCCCTGGAAGTG ATTGTCCGGGCCAACATCACAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGATGCC TCCACAGTGATCCAGTGATGGTATACTTGGACCCCATGGCTGTGGTGCGAGAAGGAGTG CCCTGGTGGGTGATCCTCTGGCTGTACTGGCTGGGCTGCTGGTGTAGCACTGCTGGTG CTGCTCCTGTGGAAGTGTGGCTTCTTCCATCGGAGCAGCCAGAGCTCATCTTTCCACC AATATCACCGGGCCTGTCTGGCTGTGCAGCCTTCAGCCATGGAAGTTGGGGGTCCAGGG ACTGTGGGGTAACT		
	ORF Start: ATG at 162		ORF Stop: TGA at 552
	SEQ ID NO: 206	130 aa	MW at 14098.0kD
NOV36j, CG56054-11 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGFLFACPLSLEETDCYRVIDDQADMQKESKE NQWLGVSVLC		
	SEQ ID NO: 207	3972 bp	
NOV36k, CG56054-12 DNA Sequence	GGAGCGGCGGGCGGGCGGGAGGGCTGGCGGGGCGAACGTCTGGGAGACGTCTGAAAGACC AACGAGACTTTGGAGACCAGAGACGCGCCTGGGGGGACCTGGGGCTTGGGGCGTGCGAGA TTTCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCC GCGACCCTTGGGGGGCCTCCGGGATTGCTACCTTTTGGCTCCCTGCTCGTGAAGTGC TCTTCTACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGCAGCAAGGAGGGCG AGCCAGGCAGCCTCTTCGGCTTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCC AGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGACGACAGGCGAATC GCACTGGAGGCCTCTTCGTTGCCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGG ACATCGACCAGGGAGCTGATATGCAAAGGAAAGCAAGGAGAACCAGTGGTTGGGAGTCA GTGTTTCGAGGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATATGAGGCAA GGCAGCGAGTGGACCAGATCTGGAGACGCGGGATGATGATTGTCGCTGCTTGTGTCTCA GCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGAC GCCCCAAGGCCATGAACAATTGGGTTCTGCCAGCAGGGCACAGCTGCCGCCCTTCTCCC CTGATAGCCACTACCTCCTCTTGGGGCCCCAGGAACCTATAATTGGAAGGGGTTGCTTT TTGTGACCAACATTGATAGCTCAGACCCCGACAGCTGGTGTATAAACTTTGGACCCTG CTGACCGGCTCCCAGGACAGCCGGAGACTTGGCCCTCAATAGTACTTAGGCTTCTCTA TTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAGAGCTGAGCTTTGTGGCTGGAGCCCCC GCGCCAACCACAAGGGTGTGTGTTATCCTGCGCAAGGACAGCGCCAGTCCGCTGGTGC CCGAGGTTATGCTGTCTGGGAGCGCCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGG CTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGGTGCCCTACTTCTTTGAGC GCCAAGAAGAGCTGGGGGTGCTGTGTATGTGTACTTGAACGAGGGGGGTACTGGGCTG GGATCTCCCTCTCCGGCTCTGCGGCTCCCTGACTCCATGTTCCGGGATCAGCCTGGCTG TCCTGGGGGACCTCAACCAAGATGGCTTCCAGATATTGAGTGGGTGCCCCCTTTGATG GTGATGGGAAAGTCTTCATCTACCATGGGAGCAGCTGGGGGTGTGCGCCAAACCTTCA AGGTGCTGTGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGGCTACTCCCTGTCAAGCAGCT TGGATATGGATGGGAACCAATACCCTGACCTGCTGGTGGGCTCCCTGGCTGACACCGCAG TGCTCTTCAGGGCCAGACCATCCTCCATGTCTCCATGAGGTCTCTATTGCTCCACGAA		

	GCATCGACCTGGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGG TCTGTTTCAGCTACATTGCAGTCCCCAGCAGCTATAGCCCTACTGTGGCCCTGGACTATG TGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGTTCCCCGTGTGACGTTCTCTGA GCCGTAACCTGGAAGAACCAAGCACCAGGCCTCGGGCACCCTGTGGCTGAAGCACCAGC ATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTC GGGCCATTGTAGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGCAGGCTC CTGGCCAGGGGCTGCCTCCAGTGGCCCCCATCTCAATGCCACCAGCCAGCACCAGC GGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGACAAGATCTGCCAGAGCAATC TGCAGCTGGTCCACGCCCGCTTCTGTACCCGGGTGAGCGACACGGAATTCACCTCTGC CCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGAAGTGGGCGAGCAGTATTGGCC TGGAGCTGATGGTCACCAACCTGCCATCGGACCCAGCCAGCCAGCCAGGCTGATGGGGATG ATGCCCATGAAGCCAGCTCCTGGTCATGCTTCTGACTCACTGCACTACTCAGGGGTCC GGGCCCTGGACCTGCGGAGAAGCCACTTGCCTGTCCAATGAGAATGCCTCCCATGTTG AGTGTGAGCTGGGAACCCCATGAAGAGAGGTGCCAGGTACCTTCTACCTCATCCTTA GCACCTCCGGGATCAGCATTGAGACCACGGAAGTGGAGGTAGAGCTGCTGTTGGCCACGA TCAGTGAGCAGGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTCATTGAGCTGCCAC TGTCCATTGCAGGAATGGCCATTCCCCAGCAACTCTTCTTCTGTTGTTGAGGGGCG AGAGAGCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGTCAAGTATGAGGTACAGGTTT CCAACCAAGGCCAGTGCCTCAGAACCCTGGGCTCTGCCTTCTCAACATCATGTGGCCTC ATGAGATTGCCAATGGGAAGTGGTGTCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGC AGGGGCTGGGCGAGAAAGGGCTTTGCTCTCCAGGCCCAACATCTCCACCTGGATGTGG ACAGTAGGGATAGGAGGCGGCGGAGCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAGC GGCAGGAGCCCAGCATGTCTGGTGGCCAGTGTCTCTGCTGAGAAGAAGAAAAACATCA CCCTGGACTGCGCCCGGGGCACGGCCAACTGTGTGGTGTTCAGCTGCCACTCTACAGCT TTGACCGCGCGGCTGTCTGCATGTCTGGGGCCGTCTCTGGAACAGCACCTTCTGGAGG AGTACTCAGCTGTGAAGTCCCTGGAAGTATTGTCCGGGCCAACATCACAGTGAAGTCTC CCATAAAGAACTTGATGCTCCGAGATGCCTCCACAGTGATCCAGTGATGGTATACCTGG ACCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGTATCCTCCTGGCTGTACTGG CTGGGCTGTCTGGTCTAGCACTGCTGGTGTCTCTGTGGAAGATGGGATTCTTCAAAC GGGCGAAGCACCCCCCGGGCGGAGGGCCCGGATGCACACCCCATCTGGCTGCTGACGGG CATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCTTCCCATTGTCTC CAGCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAG GGTAGAGTGGGCTGCTGGTGTGCATCAAGATTTGGCAGGATCGGCTTCTCAGGGGCAC AGACCTCTCCACCTCAAGAAGTCTCTCCACCCAACTTCCCCTTAGAGTGCTGTGAGAT GAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGGTGAGAAGGGCAGGGGTGTCTG ATGCAAAGGTGGGAGAAAGGATCCTAATCCCTTCTCT'CCCATTCACCTGTGTAACAG GACCCCAAGGACCTGCCTCCCGGAAGTGCCTTAACCTAGAGGTGCGGGAGGAGGTTGT GTCAGTACTCAGGCTGCTCTCTCTAGTTTCCCTCTCATCTGACCTTAGTTTGTCTGC CATCAGTCTAGTGGTTTCGTGGTTTCGTCTATTTATAAAAAATATTTGAGAACAACAAAA AAAAAAAAAA		
	ORF Start: ATG at 162		ORF Stop: TGA at 3414
	SEQ ID NO: 208	1084 aa	MW at 118234.7kD
NOV36k, CG56054-12 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVDIDQADMQKESKE NQWLGVSVRSQPGGKIVTCAHRYEARQVRDQILETRDMIGRCFVLSQDLAIRDELDDGGE WKFCGRPQGHEQFGFCQQTAAAFSPDSHYLLFGAPGTYNWKGLLFVTNIDSSDPDLV YKTLDPADRLPGPAGDLALNSYLGFIDSGLVRAEELSFVAGAPRANHKGAVVILRKD SASRLVPEVMSGERLTSGFGYSLAVADLNSDGWPDILVAPYFFERQEELGGAVVYVLN QGGHWAGISPLRLCGSPDSMFGISLAVLGDNLNQGDFPDIAVGAPFDGDKVFIYHGSGLG VVAKPSQVLEGEAVGIKSFYSLSGSLMDGNQYPDLLVGSADTAFLFRARPILHVSHE VSIAPRSIDLEQPNACAGHSVCVDRVCFYSYIAVPSSYSPTVALDYVLADTDRLRGQV PRVTFLSRNLEEPKHQASGTVLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSTYSLQTP RLRRQAPGQGLPPVAPILNAHQPSQRAEIHFLKQCGGEDKICQSNLQLVHARFCTRVSD TEFQPLPMDVDGTTALFALSGQPVIGLELMVTNLPSPDAQPADGDDAHEAQLLVMLPDS LHYSVGRALDPAEKPLCLSNENASHVECELGNPMPKRGAVTFYLLISTSGISIEETLELV ELLLATISEQELHPVSARARVFIELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKV KYEVTVSNQGSRLRTLGSFAFLNIMWPHEIANGKWLlyPMQVELEGGQGPQKGLCSRP ILHLDVDSRDRRRRELEPPEQEPGERQEPSMSWVPVSSAEKKKNITLDCARGTANCVVF		

	SCPLYSFDRRAVLHVWGRNLNSTFLEEYSAVKSLEVIVRANITVKSSIKNLMRLDASTVI PVMVYLDPMMAVVAEGVPWWVILLAVLAGLLVLALLVLLWKMGGFFKRAKHPPAGGPGCTP HPGC		
	SEQ ID NO: 209	3583 bp	
NOV361, CG56054-13 DNA Sequence	TTGGGGCGTGCAGATTTCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGG CCGGGGCTCGGAGCCGCGACCCCTTGGGGGGCCTCCGGGATTGCTACCTTTTGGCTCCC TGCTCGTCCGAAGTCTCTTCTCACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCT TGCGCAAGGAGGGCGAGCCAGGCCTCTTCGGCTTCTGTGGCCCTGCACCGGCAGT TGCAGCCCCGACCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGTG GGCAGCAGGCGAATCGCACTGGAGGCCTCTTCTGCTTGCCCGTTGAGCCTGGAGGAGACTG ACTGCTACAGAGTGGACATCGACCGAGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACC AGTGGTTGGGAGTCAGTGTTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCAC ACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGATATGATTGGTC GCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGA AGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATTTGGGTTCTGCCAGCAGGGCACAG CTGCCGCCTTCTCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAACCTATAATT GGAAGGGGTTGCTTTTTGTGACCAACATTGATAGCTCAGACCCCGACCACTGGTGTATA AAACTTTGGACCCTTGCTGACCGGCTCCCAGGACCAAGCGGAGACTTGGCCCTCAATAGCT ACTTAGGCTTCTCTATTGACTCGGGGAAAGGTCTGGTGCGTGCAGAAGAGCTGAGCTTTG TGGCTGGAGCCCCCGCGCCAACCACAAGGGTGCTGTGGTCATCTGCGCAAGGACAGCG CCAGTCGCCTGGTGCCCGAGGTTATGCTGTCTGGGGAGCGCCTGACCTCCGGCTTTGGCT ACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGGCCAGACCTGATAGTGGGTGCC CCTACTTCTTTGAGCGCAAGAAGAGCTGGGGGGTGCTGTGTATGTGTACTTGAACCAGG GGGGTCACTGGGCTGGGATCTCCCTCTCCGGCTCTGCGGCTCCCTGACTCCATGTTTCG GGATCAGCCTGGCTGTCTTGGGGACCTCAACCAAGATGGCTTTCAGATATTGCAGTGG GTGCCCCCTTTGATGGTGATGGGAAAGTCTTCATCTACCATGGGAGCAGCCTGGGGGTTG TCGCCAAACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGGCTACT CCCTGTCAAGCAGCTTGGATATGGATGGGAACCAATACCTGACCTGCTGTGGTGGGCTCCC TGGCTGACACCGCAGTGCTCTTCAGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCT CTATTGCTCCACGAAGCATCGACCTGGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCT GTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCCAGCAGCTATAGCCCTACTG TGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGTTCCCC GTGTGACGTTCTCTGAGCCGTAACCTGGAAGAACCAAGCACCAGGCTCCGGGCACCGTGT GGCTGAAGCACCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATG TCAAAGACAAGCTTCGGGCCATTGTAGTGACCTTGTCTTACAGTCTCCAGACCCCTCGGC TCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGCCCCCATCTCAATGCCACC AGCCAGCACCACAGCGGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGACAAGA TCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCCGCTTCTGTACCCGGGTGAGCGACAGG AATTCCAACCTCTGCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGTGAGTGGC AGCCAGTCATTGGCCTGGAGCTGATGGTCACCAACCTGCCATCGGACCCAGCCCAGCCCC AGGCTGATGGGGATGATGCCATGAAGCCAGCTCCTGGTCATGCTTCTGACTCACTGC ACTACTCAGGGGTCCGGGCCCTGGACCCTGCGGAGAAGCCACTCTGCCTGTCCAATGAGA ATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCAGGTACACT TCTACCTCATCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAAGTGGAGGTAGAGC TGCTGTTGGCCACGATCAGTGAGCAGGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCT TCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCCCCAGCAACTTCTTCTCTGT GTGTGGTGAGGGGCGAGAGAGCCATGCAGTCTGAGCGGATGTGGGCAGCAAGGTCAAGT ATGAGGTACCGGTTTCCAACCAAGGCCAGTCGCTCAGAACCCTGGGCTCTGCCTTCTCTCA ACATCATGTGGCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCAATGCAGGTTG AGCTGGAGGGCGGGCAGGGGCTGGGCAGAAAGGGCTTGTCTCTCCAGGCCCAACATCC TCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGGCGGAGCTGGAGCCACCTGAGCAGC AGGAGCCTGGTGAGCGGAGGAGCCAGCATGTCTGGTGGCCAGTGTCTCTGCTGCTGAGA AGAAGAAAAACATCACCTGGACTGCGCCCCGGGACAGGCCAACTGTGTGGTGTTCAGCT GCCCACTCTACAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCCGCTCTCTGGAACA GCACCTTTCTGGAGGAGTACTCAGCTGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACA TCACAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGATGCCTCCACAGTGATCCAG TGATGGTATACTTGGACCCATGGCTGTGGTGGCAGGAAGAGTGCCTGGTGGGTGCTATCC TCCTGGCTGTACTGGCTGGGCTGCTGGTGTAGCACTGTGGTGTGCTCTCTGTGGAAGT GTGGCTTCTTCCATCGGAGCAGCCAGAGCTCATCTTTCCCAACCACTATCACCGGGCCT		

	GTCTGGCTGTGCAGCCTTCAGCCATGGAAGTTGGGGGTCCAGGGACTGTGGGATGGGATT CTTCAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCC TCGGGAAGACCGACAGCAGTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAAGT GGGCAGCCCCCGGGGGAGGGCCCGGATGCACACCCCATCCTGGCTGTGACGGGCATCC CGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCCTAG
	ORF Start: ATG at 57 ORF Stop: TGA at 3411
	SEQ ID NO: 210 1118 aa MW at 121969.6kD
NOV36l, CG56054-13 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVIDIDQADMQKESKE NQWLGVSVRSQPGGKIIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGE WKFCGRPQGHQFGFCQQTAAAFSPDSHYLLFGAPGTYNWKGLLFVTNIDSSDPDQLV YKTLDPADRLPGPAGDLALNSYLGFSDSGKGLVRAEELSFVAGAPRANHKGAVVILRKD SASRLVPEVMLSGERLTSFGYSLAVADLNSDGWPDILVGPYFFERQEELGGAVVYVLN QGGHWAGISPLRLCGSPDSMFGISLAVLGDNLQDGFDPDIAVGAPFDGDGKVFYIHGSSLG VVAKPSQVLEGEAVGIKSFYSLSGSLDMQYQPDLLVGLSLADTAVLFRARPILHVSHE VSIAPRSIDLEQPNCAGGHSVCVDLRVCFYIAPVSSYSPTVALDYVLDADTDRLRGQV PRVTFLSRNLEEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYSLQTP RLRRQAPGQGLPPVAPILNAHQPSQRAEIHFLKQCGGEDKICQSNLQLVHARFCTRVSD TEFQPLPMDVDGTTALFALSGQPVIGLELMVTNLPSDPAQPPADGDDAHEAQLLVMLPDS LHYSGVRLDPAEKPLCLSNENASHVECELGNPMKRGAVTFYLILSTSGISIIETTELEV ELLLATISEQELHPVSARARVFIELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKV KYEVTVSNQGSRLRTLGS AFLNIMWPHEIANGKWLLYPMQVELEGGQGPQKGLCSPRPN ILHLDVDSRDRRRRELEPPQQEPGERQEPSMSWVPVSSAEKKKNITLDCARGTANCVFV SCPLYSFDRAAVLHVWGRWNSTFLEEYSVAKSLEVIVRANITVKSSIKNMLRDASTVI PVMVYLDPMVAEVPWWVILLAVLAGLLVLALLVLLWKCGFFHRSSQSSSFPTNYHR ACLAVQPSAMEVGGPGTGVWDSSNGRSTPRPPCPSTMR
	SEQ ID NO: 211 3938 bp
NOV36m, CG56054-14 DNA Sequence	TTGGGGCGTGCGAGATTTCCTTGCAATTCGCTGGGAGCTCGCGCAGGGATCGTCCCATG GCCGGGGCTCGGAGCCGCGACCTTGGGGGGCTCGGGGATTGCTACCTTTTGGGCTC CCTGCTCGTCGAAGTCTCTTCTCACGGGCTGTGCCTTCAATCTGGACGTGATGGGTG CCTTGCACAAGGAGGGCGAGCCAGGCAGCCTTCTCGGCTTCTGTGGCCCTGCACCGG CAGTTGCAGCCCCGACCCAGAGCTGGCTGTGTTGGTGGGTGCTCCCCAGGCCCTGGCTCT TCCTGGGCAGCAGCGAATCGCACTGGAGGCCTTTCGCTTGCCGTTGAGCCTGGAGG AGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAG GAGAACCAGTGGTTGGGAGTCACTGTTTCGAGCCAGGGGCTGGGGGCAAGATTGTTAC CTGTGCACACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATA TGATTGGTGCCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATGGT GGGGAATGGAAGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATTTGGGTTCTGCCA GCAGGGCACAGCTGCCGCTTCTCCCTGATAGCCACTACCTCCTCTTGGGGCCCTCAG GAACCTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGAC CTGGCACACCTGGACGACGCTCCCTACGAGGCGGGGGAGAGAAGGAGCAGGACCCCCG CCTCATCCCGGTCCCTGCCAACAGCTACTTTGGCTTCTCTATTGACTCGGGGAAAGGTC TGGTGGCTGCAGAAGAGCTGAGCTTTGTGGCTGGAGCCCCCGCGCAACCAAGGGT GCTGTGTTATCCTGCGCAAGGACAGCGCCAGTGCCTGGTGCCCGAGGTTATGCTGTC TGGGGAGCGCCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTG ATGGCTGGCCAGACCTGATAGTGGGTGCCCCCTACTTCTTTGAGCGCAAGAAGAGCTG GGGGGTCTGTGTATGTGTAAGTGAACAGGGGGTCACTGGGCTGGGATCTCCCTCT CCGGCTCTGCGGCTCCCTGACTCCATGTTTCGGGATCAGCCTGGCTGTCTGGGGGACC TCAACCAAGATGGCTTCCAGATATTGCAGTGGGTGCCCCCTTTGATGGTGATGGGAAA GTCTTCTATCTACCATGGGAGCAGCCTGGGGTGTGCGCAAAACCTTACAGGTGCTGGA GGGCGAGGCTGTGGGCATCAAGAGCTTCGGCTACTCCCTGTGAGGCAGCTTGATATGG ATGGGAACCAATACCTGACCTGCTGGTGGGTCCCTGGCTGACACCGCAGTCTCTTC AGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAAGCATCGA CCTGGAGCAGCCCACTGTGCTGCGGGCCACTCGGTCTGTGTGGACCTAAGGCTCTGTT TCAGCTACATTGCACTGCGGAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTA GATGCGGACACAGACCGGAGGCTCCGGGGCCAGGTTCCCGTGTGACGTTCTGAGCCG TAACCTGGAAGAACCACAGCAGGCTCGGGCACCCTGTGGCTGAAGCACCAGCATG ACCGAGTCTGTGGAGAGCCATGTTCCAGTCCAGGAAAATGTCAAAGACAAGCTTCGG

	<p>GCCATTGTAGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCC TGGCCAGGGGCTGCCTCCAGTGGCCCCATCCTCAATGCCACCAGCCAGCAGCCAGC GGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGACAAGATCTGCCAGAGCAAT CTGCAGCTGGTCCACGCCCCGCTTCTGTACCCGGGTCAGCGACACGGAATTCACACCTCT GCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGAAGTGGGAGGCTGATG GCCTGGAGCTGATGGTACCAACCTGCCATCGGACCCAGCCAGCCAGCCAGGCTGATGG GATGATGCCCATGAAGCCAGCTCCTGGTCATGCTTCTGACTCACTGCACTACTCAGG GGTCCGGGCCCTGGACCTGCGGAGAAGCCACTCTGCCTGTCCAATGAGAATGCCTCCC ATGTTGAGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCAGGTACCTTCTACCTC ATCCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAAGTGGAGGTAGAGCTGCTGTT GGCCACGATCAGTGAGCAGGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTCATTG AGCTGCCACTGTCCATTGCAGGAATGGCCATTCCCAGCAACTCTTCTCTGTTGTG GTGAGGGGCGAGAGAGCCATGCAGTCTGAGCGGGATGTGGGAGCAAGGTCAAGTATGA GGTCACGGTTTCCAACCAAGGCCAGTGCCTCAGAACCCTGGGCTCTGCCTTCTCAACA TCATGTGGCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCAATGCAGGTTGAG CTGGAGGGCGGGCAGGGGCTGGGCAGAAAGGGCTTTGCTCTCCAGGCCCAACATCCT CCACCTGGATGTGGACAGTAGGGATAGGAGCGGCGGGAGCTGGAGCCACCTGAGCAGC AGGAGCCTGGTGAAGCGCAGGAGCCAGCATGTCTGGTGGCAGTGTCTCTGCTGAG AAGAAGAAAAACATCACCTGGACTGCGCCCGGGGCAGGCCAAGTGTGTGGTGTTCAG CTGCCACTCTACAGCTTTGACCGCGGGCTGTGCTGCATGTCTGGGGCGCTCTCTGGA ACAGCACCTTTCTGGAGGAGTACTCAGCTGTGAAGTCCCTGGAAGTATTGTCCGGGCC AACATCACAGTGAAGTCTCCATAAAGAACTGATGCTCCGAGATGCCTCCACAGTGAT CCCAGTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCTGGTGGG TCATCCTCTGGCTGTACTGGCTGGGCTGTGGTGTAGCACTGTGGTGTGCTCTCTG TGGAAGATGGGATTCTTCAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCAAGTACCA TGCGGTGAAAATTCCTCGGGAAGACCGACAGCAGTTCAAGGAGGAGAAGACGGGCACCA TCCTGAGGAACAACCTGGGGCAGCCCCATCCTGGCTGGGCGCCGATGGGCATCCAGGGC CAGGCACCGCCTAGGTTCCCATGTCCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCC CCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGTGGGCTGTGGTGTGCGATCAAGAT TTGGCAGGATCGGCTCCTCAGGGCACAGACCTCTCCCCCACAAGAACTCTCCCAAC CAACTTCCCCTTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGG GTAGGGTGAGAAGGGCAGGGGTGTCTGATGCAAAGGTGGGGAGAGGGATCCTAATCC CTTCTCTCCCATTCACCTGTGTAAACAGGACCCCAAGGACCTGCCTCCCCGGAAGTGC CTTAACCTAGAGGGTCGGGGAGGAGTTGTGTCACTGACTCAGGCTGCTCCTTCTCTAG TTTCCCCTCTCATCTGACCTTAGTTTGTGCCATCAGTCTAGTGGTTTCGTGGTTTCGT CTATTTATTAATAAATATTGAGAACAAAAA</p>
	<p>ORF Start: ATG at 57</p>
	<p>ORF Stop: TGA at 3621</p>
	<p>SEQ ID NO: 212</p>
	<p>1188 aa</p>
	<p>MW at 130044.2kD</p>
<p>NOV36m, CG56054-14 Protein Sequence</p>	<p>MAGARSDPWAGSGICYLFGSLVLELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALH RQLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVIDQDQADMQKES KENQWLGVSVRSQGPQKIVTCAHYEARQVRDQILETRDMIGRCFVLSQDLAIRDELD GGEWKFCGRPQGHQFQFCQQTAAAFSPDSHYLLFGAPGTYNWKGARVELCAQGS DLAHLDDGPYEAGGEKEQDPRLLPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANH GAVVILRKDSASRLVPEVMSGERLTSFGYSLAVADLNSDGPDLIVGAPYFFERQEE LGGAVVYVLNQGHWAGISPLRLCGSPDSMFGISLAVLGLNQGDPDIAGVAPFDG KVFIYHGSSLGVVAKPSQVLEGEAVGIKSPGYSLSGLDMDGNQYDPLLVGSLADTAVL FRARPILHVSHEVSIAPRSIDLEQPNACGHSVCVDLRVCFYSIAPVSSYSPTVALDYV LDADTDRLRGQVPRVTFLSRNLEEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKL RAIVVTLSSYSLQTPRLRRQAPGQGLPPVAPILNAHQPSQRAEIHFLKQCGEDKICQS NLQLVHARFCTRVSDTEFQPLPMDVDGTTALFALSGQPVIGLELMVTNLPSPDAQPPAD GDDAHEAQLLVMLPDSLHYSVGRALDPAEKPLCLSNENASHVECELGNPMPKRAQVTFY LILSTSGIS IETTELLEVELELLATISEQELHPVSARARVFIPLSLIAGMAIPQQLFFSG VVRGERAMQSERDVGSKVKEVTVSNQGSRLTLGSAFLNIMWPHEIANGKWLlyPMQV ELEGQGPQKGLCSRPNHLHLDVDSRDRRRRELEPPEQEPGERQEPSMSWWPVSSA EKKKNITLDCARGTANCVVFSCPLYSFDRRAVLHVWGRWNSTFLEEYSVAVKSLEIVR ANITVSSSIKNMLRDASTVIVMVYLDPMVVAEGVPWWVILLAVLAGLLVLALLVL LWKMGFFKRAKHPEATVPQYHAVKI PREDRQQFKEEKTGTILRNWGSPPHGWAPMGIQ QAPPRFPCPSLACGCPPSLPQRWLLGMKRVEWAAGVASRFGRIGFLRAQTSPPTRTPP</p>

	TQLPLRLV		
	SEQ ID NO: 213	2471 bp	
NOV36n, CG56054-15 DNA Sequence	TTGGGGCGTGCAGATTTCCTTGCAATTCGCTGGGAGCTCGCGCAGGGATCGTCCCATGG CCGGGGCTCGGAGCCGCGACCCTTGGGGGGCTCCGGGATTGTGACCTTTTGGCTCCC TGCTCGTGAAGTCTCTTCTCACGGGCTGTCGCCTTCAATCTGGACGTGATGGGTGCCT TGCACAAGGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGT TGCAGCCCCGACCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCCTG GGCAGCAGGCGAATCGCACTGGAGGCCTTTCGCTTGCCCGTTGAGCCTGGAGGAGACTG ACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACC AGTGGTTGGGAGTCACTGTTTCGGAGCCAGGGGCTGGGGCAAGATTGTTACCTGTGCAC ACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTC GCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGA AGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATTGGGTTCTGCCAGCAGGGCACAG CTGCCGCCCTTCTCCCTGATAGCCACTACCTCCTCTTGGGGCCCCAGGAACCTATAATT GGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGG ACGACGGTCCCTACGAGGCGGGGGAGAGAAGGAGCAGGACCCCGCCTCATCCCGGTCC CTGCCAACAGCTACTTTGGCTTCTTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAG AGCTGAGCTTTGTGGCTGGAGCCCCCGCGCAACCAAGGGTGTCTGTGTTATCCTGC GCAAGGACAGCGCCAGTCGCCTGGTGCCGAGGTTATGCTGTCTGGGGAGCGCCTGACCT CCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGGCCAGACCTGA TAGTGGGTGCCCCCTACTTCTTTGAGCGCAAGAAGAGCTGGGGGGTGCTGTGTATGTGT ACTTGAACAGGGGGGTCACTGGGCTGGGATCTCCCTCTCCGGCTCTGCAACTCCCCGC ACTCCATGTTCCGGATCAGCCTGGCTGTCTGGGGGACCTCAACCAAGATGGCTTTCCAG ATATTGCAGTGGGTGCCCCCTTTGATGGTGTATGGGAAAGTCTTCATCTACCATGGGAGCA GCCTGGGGGTTGTGCGCAACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGA GCTTCGGCTACTCCCTGTGAGGCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGC TGGTGGGCTCCCTGGCTGACACCGCAGTGTCTTCAGGGCCAGACCCATCTCCATGTCT CCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCTGGAGCAGCCCACTGTGCTGGCG GCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCACTCCCGAGCAGT ATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGG GCCAGGTTCCCGTGTGACGTTCTGAGCCGTAACCTGGAAGAACCAAGCACCAGGCCT CCGGGACCGTGTGGCTGAAGCACCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGC TCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTGTAGTGACCTTGTCTACAGTCTCC AGACCCCTCGGCTCCGGCGGGAGGGCCGGATGCACACCCATCTGGCTGCTGAGGGC ATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCCTAGGTTCCCATGTCCC AGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGG GTAGAGTGGGCTGTGGTGTGCGATCAAGATTGGCAGGATCGGCTTCCTCAGGGGCACA GACCTCTCCACCCACAAGAACTCTCCACCCCAACTTCCCCTTAGAGTGTGTGAGATG AGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGGTGAGAAGGGCAGGGGTGTCTGA TGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTCACCTGTGTAACAGG ACCCCAAGGACCTGCCTCCCCGGAAGTGCCTTAACCTAGAGGGTCGGGGAGGAGGTTGTG TCACTGACTCAGGCTGCTCCTTCTCTAGTTTCCCTCTCATCTGACCTTAGTTTGCTGCC ATCAGTCTAGTGGTTTCGTGTTTCGTCTATTTATTAATAAATATTGAGAACAAAAAA AAAAAAAAAA		
	ORF Start: ATG at 57		ORF Stop: TAG at 1965
	SEQ ID NO: 214	636 aa	MW at 68715.7kD
NOV36n, CG56054-15 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVDDIQADMQKESKE NQWLGVSVRSQPGGKIVTCAHRYEARQVRDQILETRDMIGRCFVLSQDLAIRDELDDGE WKFCGRPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGARVELCAQGSADLAH LDDGPYEAGGEKEQDPRLLPVPANSYFGFSDSGKGLVRAEELSFVAGAPRANHKGAVVI LRKDSASRLVPEVMLSGERLTSGFGYSLAVADLNSDGPDLIVGAPYFFERQEEELGGAVY VYLNQGGHWAGISPLRLCNSPHSMFGISLAVLGDLDNQGDFPDIAVGAPFDGDKVFIYHG SSLGVVAKPSQVLEGEAVGIKSFYSLSGSLMDGNQYPDLLVGLADTAVLFRAPRILH VSHEVSIAPRSIDLEQPNACGGHSVCVDRVCFSYIAPSSYSPTVALDYVLDADTDRLR RGQVPRVTFLSRNLEEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYS LQTPRLRREGPDAPHPIAADGHPGLGPDGHPGPGTA		

	SEQ ID NO: 215	1924 bp	
NOV360, CG56054-16 DNA Sequence	TTGGGGCGTGCGAGATTTCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGG CCGGGGCTCGGAGCCGCGACCCTTGGGGGGCCTCCGGGATTGTGACCTTTTGGCTCCC TGCTCGTGAAGTGTCTTCTCACGGGCTGTCGCCTTCAATCTGGACGTGATGGGTGCCT TGGCAAGGAGGGCGAGCCAGGCAGCCTCTCGGCTTCTCTGTGGCCCTGCACCGGCAGT TGCAGCCCCGACCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTG GGCAGCAGGCGAATCGCACTGGAGGCCCTTTCGCTTGCCCGTTGAGCCTGGAGGAGACTG ACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACC AGTGGTTGGGAGTCAGTGTTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCAC ACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTC GCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATTGGTGGGGAATGGA AGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATTGGGTTCTGCCAGCAGGGCACAG CTGCCGCCCTTCTCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAACCTATAATT GGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGG ACGACGCTCCCTACGAGGCGGGGGAGAGAAGGAGCAGGACCCCGCCTCATCCCGGTCC CTGCCAACAGCTACTTTGGCTTCTTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAG AGCTGAGCTTTGTGGCTGGAGCCCCCGGCCAACCACAAGGTGCTGTGGTTCATCCTGC GCAAGGACAGCGCCAGTCGCCTGGTGCCCGAGGTTATGCTGTCTGGGGAGCGCTGACCT CCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGGCCAGACCTGA TAGTGGGTGCCCCCTACTTCTTTGAGCGCAAGAAGAGCTGGGGGGTGTGTGTATGTGT ACTTGAACAGGGGGGTCACTGGGCTGGGATCTCCCTCTCCGGCTCTGCGGCTCCCTG ACTCCATGTTTCGGGATCAGCTGGCTGTCTGGGGACCTCAACCAAGATGGCCTTCCAG ATATTGCAGTGGGTGCCCTTTGATGGTGATGGGAAAGTCTTATCTACCATGGGAGCA GCCTGGGGGTTGTGCGCAAGCCTTACAGGTGCTGGAGGGCGAGGCTGTGGGCATCCCGA GCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCCTAGGTTCCCATGTCCAGCCTGG CCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT GGGCTGCTGGTGTGCGATCAAGATTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCT CCCACCCACAAGAATCCTCCACCCAACTTCCCTTAGAGTGTGTGAGATGAGAGTGG GTAAATCAGGGACAGGGCCATGGGGTAGGGTGAGAAGGGCAGGGGTGTCTGATGCAAG GTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTACCTGTGTAACAGGACCCCAA GGACCTGCCTCCCCGGAAGTGCTTAACCTAGAGGGTCGGGGAGGAGGTTGTGTCACTGA CTCAGGCTGCTCCTTCTCTAGTTTCCCCTCTCATCTGACCTTAGTTTGCTGCCATCAGTC TAGTGGTTTCGTGGTTTCGTCTATTATTAAAAAATTTTGAAGACAAAAAATAAAAAA AAAA		
	ORF Start: ATG at 57		ORF Stop: TGA at 1671
	SEQ ID NO: 216	538 aa	MW at 57824.0kD
NOV360, CG56054-16 Protein Sequence	MAGARSRDPWASGICYLFGSLLEVLLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRQSWLLVGAPQALALPQQANRTGGLFACPLSLEETDCYRVIDDQADMQKESKE NQWLGVSVRSQGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGE WKFCGRPQGHEQFGFCQQTAAAFSPDSHYLLFGAPGTYNWKGTAARVELCAQGSADLAH LDDGPYEAGGEKEQDPRLLIPVPANSYFSGSIDSGKGLVRAEELS FVAGAPRANHKGAVVI LRKDSASRLVPEVMLSGERLTSFGYSLAVADLNSDGWPD LIGAPYFFERQELGGAVY VYLNQGGHWAGISPLRLCGSPDSMFGISLAVLGDLDLPDI AVGAPFDGDGKVF IYHG SSLGVVAKPSQVLEGEAVGIPSWAPMGIQQAAPPFPCLACGCPPSLPQRWLLGMKRV EWAAGVASRFRIGFLRGTDLSHPQELLPNFPLECCEMRVGKSGTGWPGRVRRAGVS		
	SEQ ID NO: 217	2082 bp	
NOV36p, CG56054-17 DNA Sequence	TTGGGGCGTGCGAGATTTCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGG CCGGGGCTCGGAGCCGCGACCCTTGGGGGGCCTCCGGGATTGTGACCTTTTGGCTCCC TGCTCGTGAAGTGTCTTCTCACGGGCTGTCGCCTTCAATCTGGACGTGATGGGTGCCT TGGCAAGGAGGGCGAGCCAGGCAGCCTCTCGGCTTCTCTGTGGCCCTGCACCGGCAGT TGCAGCCCCGACCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTG GGCAGCAGGCGAATCGCACTGGAGGCCCTTTCGCTTGCCCGTTGAGCCTGGAGGAGACTG ACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACC AGTGGTTGGGAGTCAGTGTTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCAC ACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTC GCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATTGGTGGGAATGGA AGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATTGGGTTCTGCCAGCAGGGCACAG		

	CTGCCGCCTTCTCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAACCTATAATT GGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGG ACGACGGTCCCTACGAGGCGGGGGAGAGAAGGAGCAGGACCCCGCCTCATCCCGGTCC CTGCCAACAGCTACTTTGGCTTCTCTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAG AGCTGAGCTTTGTGGCTGGAGCCCCCGGCCAACCAAGGGTGTGTGGTATCCTGCG GCAAGGACAGCGCCAGTCGCCTGGTGCCCGAGGTTATGCTGTCTGGGGAGCGCCTGACCT CCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGGCCAGACCTGA TAGTGGGTGCCCCCTACTTCTTTGAGCGCAAGAAGAGCTGGGGGGTGTGTGTATGTGT ACTTGAACAGGGGGGTCACTGGGCTGGGATCTCCCCCTCTCCGGCTCTGCGGCTCCCCTG ACTCCATGTTCCGGATCAGCCTGGCTGTCTGGGGGACCTCAACCAAGATGGCTGTGGTG GCAGAAGGAGTGCCTGGTGGGTATCCTCTGGCTGTACTGGCTGCTGGTGTCTGGTGCTA GCACTGCTGGTGTCTCTGTGGAAGATGGGATTCTTCAAACGGGCGAAGCACCCCGAG GCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTCGGGAAGACCGACAGCAGTTCAAG GAGGAGAAGACGGGCACCATCTGAGGAACAAGTGGGCGAGCCCCCGCGGGAGGGCCCCG GATGCACACCCCATCTGGCTGTGACGGGCATCCGAGCTGGGCCCCGATGGGCATCCA GGGCCAGGCACCGCTAGGTTCCCATGTCCCAGCCTGGCTGTGGCTGCCCTCCATCCCT TCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGTGGGCTGCTGGTGTGCGATCAAG ATTTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAACTCCTCCC ACCCAACTTCCCCTTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATG GGGTAGGGTGAGAAGGGCAGGGGTGTCTGATGCAAAGGTGGGGAGAAGGGATCCTAATC CCTTCTCTCCCATTCACCTGTGTAACAGGACCCCAAGGACCTGCTCCCGGAAGTGC CTTAACCTAGAGGGTCGGGGAGGAGTTGTGTCACTGACTCAGGCTGCTCCTTCTCTAGT TTCCCTCTCATCTGACCTTAGTTTGCTGCCATCAGTCTAGTGGTTTCGTGGTTTCGTCT ATTTATTAATAAATATTTGAGAACAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 57		ORF Stop: TGA at 1524
	SEQ ID NO: 218	489 aa	MW at 51813.5kD
NOV36p, CG56054-17 Protein Sequence	MAGARSRDPWASGICYLFGSLVLELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVDIDQADMQKESKE NQWLGVSVRSQPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGE WKFCGRPQGHQFGFCQQTAAAFSPDSHYLLFGAPGTYNWKGARVELCAQGSADLAH LDDGPYEAGGEKEQDRLIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVI LRKDSASRLVPEVMLSGERLTSGFGYSLAVADLNSDGPDLIVGAPYFFERQEELGGAVY VYLNQGGHWAGISPLRLCGSPDSMFGISLAVLGDLDNQGCGRRSALVGHPPGCTGWAAG ASTAGAAPVEDGILQTGEAPRGHRAPVPCGEDSSGRPTAVQGGEDGHHPEQLGQPPAGG PGCTPHPGC		
	SEQ ID NO: 219	3879 bp	
NOV36q, CG56054-18 DNA Sequence	TTGGGGCGTGCGAGATTCCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCCATGG CCGGGGCTCGGAGCCGCGACCTTGGGGGGCCTCCGGGATTGTCTACCTTTTGGTCCC TGCTCGTGAAGTCTCTTCTCAGGGCTGTGCGCTCAATCTGGACGTGATGGTGGCTCCT TGCGCAAGGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGT TGCAGCCCCGACCCAGAGCTGGCTGTGTGGTGGTGTCCCCAGGCCCTGGCTCTTCCTG GGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTCCCCGTTGAGCCTGGAGGAGACTG ACTGCTACAGAGTGGACATCGACAGGGAGCTGATATGCAAAGGAAAGCAAGGAGAACC AGTGGTTGGGAGTCAGTGTTCCGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCAC ACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTC GCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGA AGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATTGGGTTCTGCCAGCAGGGCACAG CTGCCGCCTTCTCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAACCTATAATT GGAAGGCGACGGCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGG ACGACGGTCCCTACGAGGCGGGGGAGAGAAGGAGCAGGACCCCGCCTCATCCCGTCC CTGCCAACAGCTACTTTGGCTTCTCTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAG AGCTGAGCTTTGTGGCTGGAGCCCCCGGCCAACCAAGGGTGTGTGGTTATCCTGCG GCAAGGACAGCGCCAGTCGCCTGGTGCCCGAGGTTATGCTGTCTGGGGAGCGCCTGACCT CCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGGCCAGACCTGA TAGTGGGTGCCCCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGGTGGTGTGTGTGT ACTTGAACAGGGGGGTCACTGGGCTGGGATCTCCCCCTCTCCGGCTCTGCGGCTCCCCTG ACTCCATGTTCCGGATCAGCCTGGCTGTCTGGGGGACCTCAACCAAGATGGCTTTCCAG ATATTGCAGTGGGTGCCCTTTGATGGTGTGGGAAAGTCTTCATCTACCATGGGAGCA		

	GCCTGGGGGTGTGCGCAAACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGA GCTTCGGCTACTCCCTGTCAGGCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGC TGGTGGGCTCCCTGGCTGACACCGCAGTGTCTTCAGGGCCAGACCCATCCTCCATGTCT CCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCTGGAGCAGCCCAACTGTGCTGGCG GCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCCCAGCAGCT ATAGCCCTACTGTGGCCCTGGACTATGTGTAGATGCGGACACAGACCGGAGGCTCCGGG GCCAGGTTCCCGTGTGACGTTCTGAGCCGTAACCTGGAAGAACCAAGCACCAGGCCT CGGGCACCGTGTGGCTGAAGCACCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGC TCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTGTAGTGACCTTGTCTACAGTCTCC AGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGTGCCTCCAGTGGCCCCATCC TCAATGCCACCAGCCAGCACCAGCGGGCAGAGATCCACTTCTGAAGCAAGGCTGTG GTGAAGACAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGG TCAGCGACACGGAATCCAACCTCTGCCATGGATGTGGATGGAAACAACAGCCCTGTTTG CACTGAGTGGGCAGCCAGTCAATTGGCCCTGGAGCTGATGGTACCAACCTGCCATCGGACC CAGCCCAGCCCCAGGCTGATGGGGATGATGCCCATGAAGCCCAGCTCCTGGTCATGCTTC CTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCTGCGGAGAAGCACTCTGCC TGTCCAATGAGAATGCCCTCCCATGTTGAGTGTGAGCTGGGGAACCCATGAAGAGAGGTG CCCAGGTCACCTTCTACCTCATCCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAAC TGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCAGGAGCTGCATCCAGTCTCTGCAC GAGCCCGTGTCTTCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCCCCAGCAAC TCTTCTTCTCTGGTGTGGTGAGGGGCGAGAGGCCATGCAGTCTGAGCGGGATGTGGCA GCAAGGTCAAGTATGAGGTACGGTTTCCAACCAAGGCCAGTCGCTCAGAACCCTGGGCT CTGCCTTCTCAACATCATGTGGCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACC CAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCTGGGCAGAAAGGGCTTGTCTCTCCA GGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGGGAGCTGGAGC CACCTGAGCAGCAGGAGCCTGGTGAGCGGCAGGAGCCAGCATGTCTGGTGGCCAGTGT CCTCTGCTGAGAAGAAGAAAACATCACCTGGACTGCGCCCCGGGCGACGCCCACTGTG TGGTGTTCAGCTGCCCCACTCTACAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCC GTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGCTGTGAAGTCCCTGGAAGTGATTG TCCGGGCCAACATCACAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGATGCCTCCA CAGTGATCCCAGTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCT GGTGGGTATCCTCCTGGCTGTACTGGCTGGGCTGTGGTGTAGCACTGCTGGTGTGCTG TCCTGTGGAAGATGGGATTCTTCAAACGGGCGAAGCACCCCCGGGCGGGAGGGCCCGGAT GCACACCCCATCCTGGCTGCTGACGGGCATCCCGAGCTGGGCCCGATGGGCATCCAGGG CCAGGCACCGCCTAGGTTCCCATGTCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCC CCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGTGGGCTGCTGGTGTGCGATCAAGATT TGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAACTCCTCCACCC CAACTTCCCCTTAGAGTGTGTGAGATGAGAGTGGGTAATCAGGGACAGGGCCATGGGG TAGGGTGAGAAGGGCAGGGGTGTCTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCT TCCTCTCCCATTCACCTGTGTAAACAGGACCCCAAGGACCTGCCTCCCCGAAGTGCCTT AACCTAGAGGGTCCGGGAGGAGGTTGTGTCACTGACTCAGGCTGCTCCTTCTCTAGTTTC CCTCTCATCTGACCTTAGTTGTCTGCCATCAGTCTAGTGGTTTCGTGGTTTCGTCTATT TATTAATAAATATTTGAGAACAAAAA		
	ORF Start: ATG at 57		ORF Stop: TGA at 3321
	SEQ ID NO: 220	1088 aa	MW at 118618.0kD
NOV36q, CG56054-18 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGGLFACPLSLEETDCYRVIDIQADMQESKE NWLGVSVRSQGPQGGKIVTCAHRYEARQQRVDQILETRDMIGRCFVLSQDLAIRDELGGGE WKFCBGRPQGHEQFGFCQQTAAAFSPDSHYLLFGAPGTYNWKG TARVELCAQGSADLAH LDDGPYEAGGEKEQDPRLI PVFANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVI LRKDSASRLVPEVMSGERLTSGFGYSLAVADLNSDGWPD LIVGAPYFFERQEELGGAVY VYLNQGGHWAGISPLRLCGSPDSMFGISLAVLGDNLNQGDFPDIAVGAPFDGDKVFIYHG SSLGVVAKPSQVLEGEAVGKISFGYSLSGSLDMDGNQYPDLLVGLADTAVLFRARPILH VSHEVSIAPRSIDLEQPNACAGHSVCVDLRVCFYSYIAVPSSYSPTVALDYLVDADTDRL RGQVPRVTFLSRNLEEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYS LQTPRLRRQAPQGGLPPVAPILNAHQPSQRAEIHFLKQCGGEDKICQSNLQLVHARFCT RVSDETFQPLPMDVDGTTALFALSGQPVIGLELMVTNLPSPDPAQPQADGDGDAHEAQLLVM LPDSLHYSGVRALDPAEKPLCLSNENASHVECELGNPMKRGAVQVTFYLILSTSGISIIET		

	ELEVELLATISEQELHPVSARARVFIELPLSIAGMAIPQQLFFPSGVVRGERAMQSERDV GSKVKYEVTVSNQGSRLRTLGSAPFLNIMWPHEIANGKWLlyPMQVELEGGQGPQKGLCS PRPNILHLDVDSRDRRRRELEPPEQQEPGERQEPMSWVPVSSAEKKKNITLDCARGTAN CVVFSCPLYSFDRAAVLHVWGRWNSTFLEEYSAVKSLEIVIRANITVKSSIKNLMRLDA STVIPVMVYLDPMVVAEGVPWWVILLAVLAGLLVALLVLLLWKMGGFFKRAKHPPAGGP GCTPHPGC		
	SEQ ID NO: 221	2709 bp	
NOV36r, CG56054-19 DNA Sequence	GGGCTTGGGGCGTGCAGATTTCCTTGCATTGCTGGGAGCTCGCGCAGGGATCGTCCC ATGGCCGGGGCTCGGAGCCGCGACCTTGGGGGGCTCCGGGATTGTACTTTTGGC TCCCTGCTCGTGAAGTCTCTTCTCACGGGCTGTGCGCTTCAATCTGGACGTGATGGGT GCCTTGCACAAGGAGGGCGAGCCAGGCAGCCTCTTGGCTTCTCTGTGGCCCTGCACCGG CAGTTGCAGCCCCGACCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTT CCTGGGCAGCAGCGAATCGCACTGGAGGCTCTTGGCTTGGCCCTTGGCCTGGAGGAG ACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGAAAAGGAAAGCAAGGAG AACCAGTGGTTGGGAGTCACTGTTTGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGT GCACACCGATATGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATT GGTCGCTGCTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGGAA TGGAACTTCTGTGAGGGACGCCCCAAGGCCATGAACAATTGGGTTCTGCGCAGCGGGC ACAGCTGCCGCCTTCTCCCCTGATAGCCACTACCTCTTGGGGCCCCAGGAACCTAT AATTGGAAGGGGTGCTTTTTGTGACCAACATTGATAGCTCAGACCCGACAGCTGGTG TATAAACTTTGGACCCTGCTGACCGGCTCCAGGACAGCCGGAGACTTGGCCCTCAAT AGCTACTTAGGCTTCTCTATTGACTCGGGGAAAGGTCTGGTGGTGCAGAGAGCTGAGC TTTGTGGCTGGAGCCCCCGCGCAACCAAGGGTGTGTGGTCACTTCCGCGCAAGGAC AGCGCCAGTCGCTGGTGGCCGAGGTTATGCTGTCTGGGAGCGCCTGACCTCCGGCTTT GGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGTGGCCAGACCTGATAGTGGGT GCCCCCTACTTCTTGGAGCGCAAGAAGAGCTGGGGGGTGTGTGTATGTGTACTTGAAC CAGGGGGGTCACTGGGCTGGGATCTCCCTCTCCGGCTCTGCGGCTCCCTGACTCCATG TTGGGATCAGCTGGCTGTCTGGGGGACCTCAACCAAGATGGCTTTCCAGATATGCA GTGGGTGCCCCCTTGTATGGTGTATGGGAAAGTCTTCATCTACCATGGGAGCAGCTGGGG GTTGTGCGCAAACTTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGGC TACTCCCTGTCAGGCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGCTGTGTGGC TCCCTGGCTGACACCGCAGTGCTCTTACGGGCCAGACCCATCCTCCATGTCTCCCATGAG GTCTCTATTGCTCCACGAAGCATCGACCTGGAGCAGCCCAACTGTGCTGGCGGCCACTCG GTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCCCAGCAGCTATAGCCCT ACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGAGGCTCCGGGGCCAGGTT CCCCGTGTGACGTTCTGAGCCGTAACCTGGAAGAACCAAGCACCAGGCCCTCGGGCACC GTGTGGCTGAAGCACCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAA AATGTCAAAGACAAGCTTCGGGCCATTGTAGTGACCTTGCTCTACAGTCTCCAGACCCCT CGGCTCCGGGCACAGGCTCTGGCCAGGGGCTGCCTCCAGGGCCTGGGCAAGAGGCTT TGCTCTCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGG GAGCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAGCGGCAGGAGCCAGCATGTCTGG TGGCCAGTGTCTCTGCTGAGAAGAAAGAAAACATCACCTGGACTGCGCCCGGGGCACG GCCAACTGTGTGGTGTTCAGCTGCCCACTCTACAGCTTTGACCGCGGGCTGTGCTGCAT GTCTGGGGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGCTGTGAAGTCCCTG GAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCTCCATAAAGAAGTGTGCTCCGA GATGCCTCCACAGTGATCCAGTGATGGTATACTTGGACCCATGGCTGTGGTGGCAGAA GGAGTGCCCTGGTGGGTGATCCTCCTGGCTGTACTGGCTGGGCTGCTGGTGTAGCACTG CTGGTGTGCTCCTGTGGAAGATGGGATTCTTCAAACGGGCGAAGCACCAGGAGCCACC GTGCCCCAGTACCATGCGGTGAAGATTCTCGGGAAGACCGACAGTTCAGAGGAGGAG AAGACGGGCACCATCCTGAGGAACAACCTGGGGCAGCCCCCGGGGAGGGCCCGGATGCA CACCCCATCCTGGCTGCTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCA GGCACCCTAGGTTCCCATGTCCAGCCTGGCTGTGGCTGCCCTCCATCCCTTCCCCA GAGATGGCT		
	ORF Start: ATG at 61		ORF Stop: TAG at 2650
	SEQ ID NO: 222	863 aa	MW at 94348.4kD
NOV36r, CG56054-19	MAGARSRDWPWASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAQALALPGQQANRTGGLFACPLSLEETDCYRVVIDQADMQKESKE NQWLGVSVRSQGGPKIVTCAHRYEARQVRDQILETRDMIGRCFVLSQDLAIRDELDDGGE		

Protein Sequence	WKFCGRPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGLLFVTNIDSSDPDQLV YKTLDPADRLPGPAGDLALNSYLGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKD SASRLVPEVMLSGERLTSGFGYSLAVADLNSDGWPD LIVGAPYFFERQEELGGAVYVYLN QGGHWAGISPLRLCGSPDSMFGISLAVLGDNLQDGFDPDIAGAPFDGDGKVFYIYHGSSLG VVAKPSQVLEGEAVGIKSFYSLSGSLMDGNQYPDLLVGS LADTAVLFRARPILHVSHE VSIAPRSIDLEQPNACAGHSVCVDLRVCFYIAVPSSYSPTVALDYVLADTDRLRGQV PRVTFLSRNLEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYLSQTP RLRRQAPGQGLPPGPGQKGLCSRPNHLHLDVDSRDRRRRELEPPEQQEPGERQEPSMSW WPVSSAEKKKNITLDCARGTANCVFSCPLYSFDRAAVLHVWRLWNSTFLEEYSVAKSL EVIVRANITVKS SIKNMLRDASTVIPVMVYLDPMVAE GVPWWVILLAVLAGLLVLAL LVLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTG TILRNWGS PRREGPDA HPILAADGHPGLGPDGHPGPTA
	SEQ ID NO: 223 4031 bp
NOV36s, CG56054-02 DNA Sequence	GGAGCGCGGGCGGGCGGGAGGGCTGGCGGGGCGAACGCTCTGGGAGACGCTCGAAAGACC AACGAGACTTTGGAGACCAGAGACGCGCCTGGGGGGACCTGGGGCTTGGGGCGTGCAGAG TTTCCCTTGCACTTCGCTGGGAGCTCGCGCAGGGATCGTCCCATGGCCGGGGCTCGGAGCC GCGACCTTGGGGGGCCTCCGGGATTGCTACCTTTTGGCTCCCTGCTCGTCAACTGC TCTTCTACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCTTGGCGAAGGAGGGCG AGCCAGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGAGTTGCAGCCCCGACCC AGAGCTGGCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTCTGGGCAGCAGGGCAATC GCACTGGAGGCCTCTTCGCTTGGCCGTTGAGCCTGGAGGAGACTGACTGTACAGAGTGG ACATCGACAGGGAGCTGATATGCAAAAGGAAAGCAAGGAGAACAGTGGTTGGGAGTCA GTGTTGCGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATATGAGGCAA GGCAGCAGTGGACCATCTTGAGAGCGCGGATATGATTGGTCTGCTGCTTGTGCTCA GCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGAATGGAAGTTCTGTGAGGGAC GCCCCAAGGCCATGAACAATTTGGGTTCTGCCAGCAGGGCAGACTGCCGCTTCTCCC CTGATAGCCACTACCTCCTCTTGGGGCCCCAGGAACCTATAATTGGAAGGGGTTGCTTT TTGTGACCAACATTGATAGCTCAGACCCGACAGCTGGTGTATAAACTTTGGACCTGT CTGACCGGCTCCAGGACACGCGGAGACTTGGCCCTCAATAGCTACTTAGGCTTCTCTA TTGACTCGGGGAAAGGTCTGGTGGTGCAGAGAGCTGAGCTTTGTGGCTGGAGCCCCC GCGCCAACCACAAGGTGCTGTGTTATCCTGCGCAAGGACAGCGCCAGTGCCTGGTGC CCGAGGTTATGCTGTCTGGGAGCGCCTGACCTCCGGCTTTGGCTACTACTGGCTGTGG CTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGGTGCCCCCTACTTCTTTGAGC GCCAAGAAGAGCTGGGGGGTCTGTGTATGTGTA CTGAAACAGGGGGTCACTGGGCTG GGATCTCCCTCTCCGGCTCTGCGGCTCCCTGACTCCATGTTGGGATCAGCCTGGCTG TCCTGGGGACCTCAACCAAGATGGCTTTCCAGATATTGCACTGGGTGCCCCCTTTGATG GTGATGGGAAAGTCTTCTATCTACCATGGGAGCAGCCTGGGGGTTGTGCGCAACCTTCAC AGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGGCTACTCCCTGTCAAGCAGCT TGGATATGGATGGGAACCAATACCTGACCTGCTGGTGGGTCCCTGGCTGACACCCGAG TGCTCTTCAGGGCCAGACCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAA GCATCGACCTGGAGCAGCCCACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGG TCTGTTTCAGCTACATTGCACTCCCGAGCAGCTATAGCCCTACTGTGGCCCTGGACTATG TGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGTTCCCCGTGTGACGTTCTCTGA GCCGTAACCTGGAAGAACCACAGCAGGCTCGGGCACCCTGTGGGTGAGACCCAGC ATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTC GGGCCATTGTAGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGGCGACAGGCTC CTGGCCAGGGGCTGCCTCCAGTGGCCCCATCCTCAATGCCACAGCCAGCAGCCAGC GGGCGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGACAAGATCTGCCAGAGCAATC TGCACTGGTCCACGCCCGCTTCTGTACCGGGTCAAGCAGACCGAATTCACCTCTGC CCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGAAGTGGGCGAGGCTCATTGGCC TGGAGCTGATGGTCAACCACTGCCATCGGACCCAGCCAGCCAGGCTGATGGGGATG ATGCCCATGAAGCCAGCTCCTGGTCATGCTTCTGACTACTGCACTACTCAGGGGTCC GGGCCCTGGACCTGCGGAGAAGCACTCTGCCTGTCCAATGAGAATGCTTCCATGTTG AGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCAGGTCACTTCTACCTCATCCTTA GCACCTCCGGGATCAGCAATTGAGACCAGGAAGTGGAGGTAGAGCTGCTGTTGGCCACGA TCAGTGAGCAGGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTCAATGAGCTGCAC TGTCATTGCAGGAATGGCCATTCCCGAGCAACTCTTCTCTCTGGTGTGGTGGGGGCG AGAGAGCCATGCAGTCTGAGCGGGATGTGGGAGCAAGGTCAAGTATGAGGTACCGGTTT CCAACCAAGGCCAGTCGCTCAGAACCTGGGCTCTGCCTTCTCAACATCATGTGGCCTC

	ATGAGATTGCCAATGGGAAGTGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGC AGGGGCCTGGGCAGAAAGGGCTTTGCTCTCCCAGGCCCAACATCCTCCACCTGGATGTGG ACAGTAGGGATAGGAGGCGGCGGGAGCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAGC GGCAGGAGCCCAGCATGTCTGGTGGCCAGTGTCTCTGCTGAGAAGAAGAAAAACATCA CCTTGGACTGCGCCCGGGGCACGGCCAACTGTGTGGTGTTCAGCTGCCCACTCTACAGCT TTGACCGCGCGGCTGTCTGCATGTCTGGGGCCGTCTCTGGAACAGCACCTTTCTGGAGG AGTACTCAGCTGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCTT CCATAAAGAACTTGATGCTCCGAGATGCCTCCACAGTGATCCCAGTGATGGTATACTTGG ACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGTATCCTCTCTGGCTGTACTGG CTGGGCTGCTGGTGTAGCACTGCTGGTGTCTCTGTGGAAGATGGGATTCTTCAAAC GGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAAAATCCTCGGGAAG ACCGACAGCAGTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACCTGGGGCAGCC CCCATCCTGGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCCTAGGTTCCCATGTC CCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGA GGGTAGAGTGGGCTGCTGGTGTGCGATCAAGATTTGGCAGGATCGGCTTCTCAGGGCAC AGACCTCTCCCCCACAAGAACTCCTCCACCCAACTTCCCCTTAGAGTGTGTGAGATG AGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGGTGAGAAGGGCAGGGGTGTCTGA TGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTACCTGTGTAACAGG ACCCCAGGACCTGCCTCCCCGGAAGTGCTTAACCTAGAGGGTCGGGGAGGAGGTTGTG TCACTGACTCAGGCTGCTCCTTCTCTAGTTTCCCCTCTCATCTGACCTTAGTTTGTGCC ATCAGTCTAGTGGTTTCGTGGTTTCGTCTATTTATTAAAAAATTTGAGAACAAAAAA AAAAAAAAAA		
	ORF Start: ATG at 162		ORF Stop: TGA at 3714
	SEQ ID NO: 224	1184 aa	MW at 129660.8kD
NOV36s, CG56054-02 Protein Sequence	MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHR QLQPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVDIQGADMQESKE NQWLGVSVRSQPGGKIVTCAHRYEARQVRDQILETRDMIGRCFVLSQDLAIRDELGGGE WKFCGRPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGLLFVTNIDSSDPQLV YKTLDPADRLPGPAGDLALNSYLGFIDSGLVRAEELSFVAGAPRANHKGAVVILRKD SASRLVPEVMSLGERLTSGFGYSLAVADLNSDGWPDIVGAPYFFERQBELGGAVVYLN QGGHWAGISPLRLCGSPDSMFGISLAVLGDNLQDGFPIAVGAPFDGDKVFIYHGSSLG VVAKPSQVLEGEAVGIKSGYSLSGSLDMDGNQYPDLLVGLADTAVLFRARPILHVSHE VSIAPRSIDLEQPNPCAGGHSVCVDLRCFSYIAVPSSYSPTVALDYVLADTDRLRGQV PRVTFLSRNLEEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYLQTP RLRRQAPGQGLPPVAPILNAHQPSQRAEIHFLKQCGEDKICQSNLQLVHARFCTRVSD TEFQPLPMDVDGTTALFALSGQPVIGLELMVTNLPSPDPAQPDAGDDAHEAQLLVMLPDS LHYSGVRALDPAEKPLCLSNENASHVECELGNPMKRGAQVTFYLILSTSGISIIETTELEV ELLLATISEQELHPVSARARVFIELPLSIAGMAIPQQLFFSGVVVRGERAMQSERDVGSKV KYEVTVSNQGSRLRTLGS AFLNIMWPHEIANGKWLlyPMQVELEGGQGPQKGLCSRPNP ILHLDVDSRDRRRRELEPPEQQEPGERQEPMSMSWVPVSSAEKKKNTLDCARGTANCVVF SCPLYSFDRAAVLHVWGRWLNSTFLEEYSVAVKSLEIVRANITVKSSIKNLMRLDASTVI PVMVYLDPMVAEAGVPWWVILLAVLAGLLVLALLVLLWKMGFFKRAKHPEATVPQYHA VKIPREDRQQFKEEKTGTILRNWGSPPHGWAPMGIQGPAPRFPCLSLACGCPPSLPQR WLLGMKRVEWAAGVASRFRIGFLRAQTSPPTRTPPTQLPLRVL		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 36B.

Table 36B. Comparison of NOV36a against NOV36b through NOV36s.		
Protein Sequence	NOV36a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV36b	1..607 1..607	590/607 (97%) 591/607 (97%)

NOV36c	1..439 1..439	423/439 (96%) 424/439 (96%)
NOV36d	1..142 1..142	129/142 (90%) 129/142 (90%)
NOV36e	1..64 1..64	64/64 (100%) 64/64 (100%)
NOV36f	1..113 1..112	84/113 (74%) 86/113 (75%)
NOV36g	1..395 1..395	382/395 (96%) 382/395 (96%)
NOV36h	1..276 1..283	225/286 (78%) 233/286 (80%)
NOV36i	1..1079 1..1083	963/1089 (88%) 969/1089 (88%)
NOV36j	1..128 1..128	115/128 (89%) 115/128 (89%)
NOV36k	1..1076 1..1076	997/1076 (92%) 997/1076 (92%)
NOV36l	1..1079 1..1079	991/1079 (91%) 993/1079 (91%)
NOV36m	1..1137 1..1134	1011/1147 (88%) 1015/1147 (88%)
NOV36n	1..607 1..611	562/617 (91%) 567/617 (91%)
NOV36o	1..439 1..443	395/449 (87%) 400/449 (88%)
NOV36p	1..395 1..399	354/405 (87%) 358/405 (87%)
NOV36q	1..1076 1..1080	969/1086 (89%) 973/1086 (89%)
NOV36r	1..606 1..606	593/606 (97%) 593/606 (97%)
NOV36s	1..1137 1..1130	1039/1137 (91%) 1039/1137 (91%)

Further analysis of the NOV36a protein yielded the following properties shown in Table 36C.

Table 36C. Protein Sequence Properties NOV36a	
PSort analysis:	0.4600 probability located in plasma membrane; 0.1363 probability located in microbody (peroxisome); 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)

SignalP analysis:	Cleavage site between residues 34 and 35
-------------------	--

A search of the NOV36a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 36D.

Table 36D. Geneseq Results for NOV36a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV36a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB36936	Integrin alpha chain 7 - <i>Homo sapiens</i> , 1137 aa. [WO200066628-A1, 09- NOV-2000]	1..1137 1..1137	1137/1137 (100%) 1137/1137 (100%)	0.0
AAU29083	Human PRO polypeptide sequence #60 - <i>Homo sapiens</i> , 1141 aa. [WO200168848-A2, 20- SEP-2001]	1..1137 1..1141	1109/1147 (96%) 1113/1147 (96%)	0.0
AAB44308	Human PRO768 (UNQ406) protein sequence SEQ ID NO:437 - <i>Homo sapiens</i> , 1141 aa. [WO200053756- A2, 14-SEP-2000]	1..1137 1..1141	1109/1147 (96%) 1113/1147 (96%)	0.0
AA Y41752	Human PRO768 protein sequence - <i>Homo sapiens</i> , 1141 aa. [WO9946281-A2, 16-SEP-1999]	1..1137 1..1141	1109/1147 (96%) 1113/1147 (96%)	0.0
AAB94058	Human protein sequence SEQ ID NO:14232 - <i>Homo sapiens</i> , 973 aa. [EP1074617-A2, 07-FEB- 2001]	159..1137 1..973	970/979 (99%) 971/979 (99%)	0.0

- In a BLAST search of public sequence databases, the NOV36a protein was found to
- 5 have homology to the proteins shown in the BLASTP data in Table 36E.

Table 36E. Public BLASTP Results for NOV36a				
Protein Accession Number	Protein/Organism/Length	NOV36a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
JC5950	integrin alpha-7 chain precursor - human, 1137 aa.	1..1137 1..1137	1137/1137 (100%) 1137/1137 (100%)	0.0

Q13683	Integrin alpha-7 precursor - <i>Homo sapiens</i> (Human), 1181 aa.	1..1137 1..1181	1137/1181 (96%) 1137/1181 (96%)	0.0
I61186	alpha-7 integrin - mouse, 1135 aa.	14..1137 14..1135	985/1124 (87%) 1046/1124 (92%)	0.0
Q61738	Integrin alpha-7 precursor - <i>Mus musculus</i> (Mouse), 1179 aa.	14..1137 14..1179	985/1168 (84%) 1046/1168 (89%)	0.0
Q63258	Integrin alpha-7 (H36-alpha7) - <i>Rattus norvegicus</i> (Rat), 1106 aa.	34..1137 1..1106	922/1110 (83%) 981/1110 (88%)	0.0

PFam analysis predicts that the NOV36a protein contains the domains shown in Table 36F.

Table 36F. Domain Analysis of NOV36a			
Pfam Domain	NOV36a Match Region	Identities/ Similarities for the Matched Region	Expect Value
FG-GAP	49..114	20/67 (30%) 48/67 (72%)	6.1e-11
FG-GAP	260..317	20/66 (30%) 42/66 (64%)	5.4e-06
FG-GAP	318..377	26/65 (40%) 49/65 (75%)	1.3e-14
FG-GAP	378..435	30/67 (45%) 51/67 (76%)	2.2e-18
FG-GAP	436..489	20/66 (30%) 42/66 (64%)	6.1e-08
integrin_A	1061..1075	7/15 (47%) 14/15 (93%)	0.0074

Example 37.

The NOV37 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 37A.

Table 37A. NOV37 Sequence Analysis		
	SEQ ID NO: 225	4096 bp
NOV37a, CG88634-01 DNA Sequence	ATCTGTTTATTATTCTGTTAATTCCAATAGTATAATTTGACATGCATTCTGTTTT GTCTTTTCAGGTGCCATTGGATTGACTTTAGTGGCACGATGTA CTCTGAGTGGAGGTC ACTGCATTGGTGATT CAGAATGATCAAGGCCATACCA GTGTGCTGCACAGCTATCCAGA GAGCGTTGGACGAGAGGTGGCAAATGCTGTAGTCCGTCCTCTGGGCAGGTGTTAGGTAC CCCTTCAGTGGCTGGTAGTGAGAATTGTTAAAACTGACAAAGAAGTAAAATGGACCAT	

GGAAGTAATTTGCTATGGACTGACCCCTCCATTGGATGGAGAGACTGTAAATATTGCGT
TGATGTATATACAGACTGGATTATGGCTTTAGTGTGCCAAAAGATTCTATTCCATTGCC
AGTTATTAAAGAGCCTAATCAATATGTTCAAACATACTAAACACCTACAGAATCTTTT
TGTACCAAGACAGGAACAGGGTTCAGTCAGATTGACTATGCTTACAGGTCTGAGAGC
CATTAGAAAACCTGGCCCGTGAGTCATCTCTCATGGCCCCGAGAACTTGGGAAGTCTTACT
GTTGTTTTCTTCTGCAGATTAACGACATACTTCTGGCCCCACCAACTGTTCAAGGTTTGAT
TGCTGAGAATCTAGCAGAGAAGTTGATTGGTGTCTCTTTGAGGTGTGGTTACTAGCTTG
TACTCGGTGCTTCCCAACACCTCCTTATTGGAAAACAGCCAAGGAGATGGTGGCTAACTG
GAGGCATCACCCAGCAGTGGTGGAGCAGTGGAGCAAGGTCATTGTGCACTCACTTCCAG
GTTACTACGCTTTACATATGGTCCTTCATTTCTGCATTTAAAGTTCCTCGATGAAGATGC
CAGTCTGATCCCTCCAGAAATGGATAATGAGTGTGTGTCACAGACATGGTTTCGCTTTTT
ACACATGTTAAGTAATCCTGTGGATTGAGTAACCCAGCTATTATAAGCTCTACTCCCAA
ATTTAGGAACAGTCTTGAATGTGAGCGGAATGCCGCAAGAAATGAATCAGTATCCCTG
CCTTAAACATCTGCCTCAAATATTTTTCTGCGCATGCGTGGAAATCAGCTGTCTGGTGGGA
TGCAATCTTAGGTATTTCTAGACCCGATCAGACAGTGTCTCCCCAACACCCGTGAATAG
ATTAAGTATGCCTCAAAGTGCTGCTGTGAGTACCAACCCCCACATAACCGGAGGCACCG
GGCTGTTACTGTGAATAAGGCCACCATGAAGACAAGCACAGTTAGTACTGCTCATGCCTC
TAAAGTTCAGCACCAGACGTCTCCACCTCTCCTCTGTCAAGTCAAATCAGACTAGTTC
AGAACCCCGGCCACTGCCTGCCCTCGGAGACCAAGGTTAACAGCATCTTGAATCTCTT
TGGATCATGGTTATTTGATGCAGCATTGTTATGGAGTTTCGACGGAAGGGTCACAAAT
GTCCACAGACACCATGGTTTTCAATCCTATGTTTTGATGCAAGTGAATTTCTGATAACTA
TGAAGCAGGAAGAGCTGAGGCTGTGGGACACTGTGTAGGATTTTTGTAGCAAGAAGAC
TGGAGAAGAGATTCTGCCAGCTATTTATCCAGATTTACATGCTTTTAAATCAAGGTTT
GCAGATAAATGATTATGTGTGCCATCCTGTCTTGCCAGCGTTATTCTAAACTCTCCTCC
TTTGTCTGCTGTGACTTGAAAGGGATTGATGTTGTGGTTCTTACTTTATTTCACTCT
TGAAACCATTTTGCTGACAGGAGAGAATCTCAAATTCAAAAGCTATGTAAATCCAAAC
AGAATTGCGAAGATCCTCCATTAATATCCTGCTTTCTTTGTTGCCCTCCCTCATCATTT
TGGCACAGTCAAATCTGAGTCTTATGATAAACCAATAACTTTCTGTCCCTGAAGTTGAG
ACTTGTGAATATATTAATAGGTGCCTTGCAAACCTGAAACGGACCCCAACACACCCAAAT
GATATTAGGTGATTGAGTGTGGGCTCCTGATTGCGAGCATTCATCTCGTCAACCAAG
ACTCAACTCCCAGTGGCGCCAAGACATGAGCATATCACTGGCAGCTCTAGAGCTCCTCTC
TGGCCTTGCAAAGGTGAGGAAGACAGACTCAGGAGACCGGAAGCGAGCCATCAGTTCTGT
GTGCACCTACATTGTTTATCAGTGTAGTCGGCCAGCTCCTTTACACTCCAGGGATCTGCA
CTCCATGATAGTGGCAGCTTTTCAGTGTCTCTGTGCTGCTGGCTGACAGAGACCCCTGATAT
GCTTGATGAAAAGGACTGCCTTAAGGAAGTACTGGAGATTGTGGAACGGGTATCTCAGG
AAGTAAGTCCAAGAACAATGAGCAAGAGGTCAAGTACAAAGGAGATAAGGAGCCAAACCC
TGCATCTATGAGGTAAGGATGCTGTGTAAGCCACCCTAACATCCATTCTCCATAGCAT
TGGCGCATTTCTTCACTAGTGGTCTGCCTCTCCTTGTAGTCTTGTGAATGAGACCAC
TTTGATTAAATACTCCAGGCTGCCAACCATAAACCAAGCATAGTTTCCGGTACTTTGTCTT
GGATAACAGTGTCTCCTGGCAATGTGGAACAACCTCTTGGAAATGAGCAGAATGATTT
TTTCCCTCTGTCACTGTGCTGGTCCGGGAATGTCTGGAAGACTTGTGGGCACAACA
GCTTGTCTTTTACCCAGAGGAGCAAAAGCAAATCAGAAGCTTTTTGTACCTGAACCTCG
CCCAGTTCCTAAAAATGACGTTGGATTAAATATTCTGTGAAACATCGGCCATTTCTGA
AGAGGTGGACAAGATTCTTTGTGAAAGCAGATCTCAGCATTCAGATTGTCATGAAAT
AGTCACTGAAGAATTAGAAGAGAGACACGAAAAATTAAGGAGTGGCATGGCCAGCAGAT
TGCTTATGAAATACACCTTGAGCAACAGAGTGAGGAGGAATTGCAGAAGAGAAGTTTTCC
TGACCCAGTTACGGATTGCAAGCCCCCGCTCCTGCCAGGAATTCAAACAGCCCGCT
TTTTCTCTCACACTTTGGATTTTTGTCTTAGAAGCACTGAAGGAACCTGCAAATAGTCG
TCTACCTCCTCACCTTATGCACTTGATTCACGATACCTGGATTTTTGATGACATTGG
GTATCTGGATCTCTTGCCATGTGCTCCTTTTGACACAGTTTTTATTTTCTATATGAAGCC
AGGTCAGAAAACGAACCAAGAGATTTAAAGAATGTGGAGTCTTCAGAACTGTTCAAGCC
ACATTTCTAGAAATTTTGTCTTCCCTTGGCTGGTCACTAGATGTGGGCAGACACCCCTGG
TTGGAAGTGGCATGTTTCTACAGTTGGTCTATTAATTGTTGTGATGATGGTGAAGGATC
TCAACAAGAAGTGATTTCTCTGAAGATATTGGAGCTAGCATTTTCAATGGACAGAAGAA
GGTGTGTATTATGCTGATGCCCTTACAGAAATGCTTTTGTGGTCTCTCTCTGTGGA
GTCCTTAACTGATTCAATGGAAAGTAACATCTCGGACCAAGATAGTGATTCAAATATGGA
TCTTATGCCAGGAATTCGAAACAGCCATCCCTGACACTTGAGCTTTTCCCAATCATAC
AGACAATCTTAATTCCTCACAGAGGCTCAGTCCAGTTCCAGAATGAGGAAGCTGCCTCA
GGGTGCGCCTGTTCTCTCCCTTGACCTGAGACAAGAGTTTCTGTAGTCTGGGTGGAACG
CTATGATGATATAGAAAATTTCCCTCTCAGAGCTGATGACAGAGATCAGTACTGGTGT

	GGAAACTACTGCAAAATAGTAGCACTTCACTGAGATCTACAACCTTGAAAAAGAAGTTCC TGTCACTCTTCATCCACCCTTTAAACACTGGATTATTCGGATAAAAATTCAAGGAGCCAC TGGAAAATTTAATATGGTCATCCCTCTTGTGGATGGGATGATTGTCAGCAGCGAGCTCT TGGCTTTCTGGTGAGG		
	ORF Start: ATG at 101		ORF Stop: end of sequence
	SEQ ID NO: 226	1332 aa	MW at 149066.8kD
NOV37a, CG88634-01 Protein Sequence	MYSEWRSLHLVIQNDQGHTSVLHSPESVGREVANAVVRPLGQVLGTPSVAGSENLLKTD KEVKWTMEVICYGLTLPDGETVKYCVDVYTDWIMALVLPKDSIPLPVIKEPNQYVQTTIL KHLQNLFPVRQEQGSSQIRLCLOVLRAIQKLARESSLMARETWEVLLLFLQINDILLAP PTVQGLIAENLAEKLIGVLFVWLLACTRCFPTPPYWKTAKEVMANWRHHPAVVEQWSKV ICALTSRLLRFTYGPSFPAPKVPDEADSLIPPEMDNECVAQTWFRFLHMLSNPVDLSNPA IISSTPKFQEQFLNVSGMPQELNQYPCLKHLPIFFRAMRGISCLVDAFLGISRPRSDSA PPTPVNRLSMPQSAAVSTTPPHNRRHRAVTVNKATMKTSTVSTAHASKVQHQSSTSPLS SPNQTSSSEPRPLPAPRRPKVNSILNLFGSWLFDAAFVMEFRRKGSQMSTDTMVSNPMFDA SEFPDNYEAGRAECGTLCRIFCSKKTGEEILPAYLSRFYMLLIQGLQINDYVCHPVLAS VILNSPPLFCDDLKIGIDVVVPYFISALETILPDRRELSKFKSYVNPTELRRSSINILLSL LPLPHHFGTVKSESYDKPITFLSLKRLNVNIGALQTTETDPNNTQMILGDSAAGLLIRS IHLVTQRLNSQWRQDMSISLALELLSGLAKVRKTDSGDRKRAISSVCTYIVYQCSRPAE LHSRDLHSMIVAFAQCCLCVWLTEHPDMLDEKDCLEIVELGISGSKSNNEQEVKYK GDKEPNPASMVRKDAAEATLTSILHSIGAFPSPSGPASPCSLVNETTLIKYSRLPTINKH SFRYFVLDNSVILAMLEQPLGNEQNDFFPSVTVLVRGMSGRLAWAQQCLLPRGAKANQK LFVPEPRPVPKNDVGFKYSVKHRPFPEEVDKIPFVKADLSIPDLHEIVTEELEERHEKLR SGMAQQIAYEIHLEQQSEBELQKRSFPDPVTDCKPPPPAQEFQTARLFLSHFGFLSLEAL KEPANSRLPPHLIALDSTIPGFFDDIGYLDLLPCRPFDTVFIYMKPGQKTNQEILKNVE SSRTVQPHFLEFLSLGWSVDVGRHPGWTGHVSTWSINCCDDGEGSQEIVISSEDIGAS IFNGQKKVLYYADALTEIAFVVPSPVESLTDLSLESNISDQSDSDSNMDLMPGILKQPSLTL ELFPNHTDNLNSSQRLSPSSRMKRLPQGRPVPLGPETRVSVVWVERYDDIENFPLSELM TEISTGVETTANSSTSLRSTTLEKEVPVIFIHPLNTGLFRIKIQGATGKFMNVIPLVDGM IVSRRALGFLVR		

Two polymorphic variants of NOV37a have been identified and are shown in Table 410. Further analysis of the NOV37a protein yielded the following properties shown in Table 37B.

Table 37B. Protein Sequence Properties NOV37a	
PSort analysis:	0.7900 probability located in plasma membrane; 0.3500 probability located in nucleus; 0.3000 probability located in microbody (peroxisome); 0.3000 probability located in Golgi body
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV37a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 37C.

Table 37C. Geneseq Results for NOV37a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV37a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value

AAM39605	Human polypeptide SEQ ID NO 2750 - <i>Homo sapiens</i> , 515 aa. [WO200153312-A1, 26-JUL-2001]	878..1332 1..456	455/456 (99%) 455/456 (99%)	0.0
AAM41391	Human polypeptide SEQ ID NO 6322 - <i>Homo sapiens</i> , 321 aa. [WO200153312-A1, 26-JUL-2001]	1072..1332 1..262	261/262 (99%) 261/262 (99%)	e-147
ABB58732	<i>Drosophila melanogaster</i> polypeptide SEQ ID NO 2988 - <i>Drosophila melanogaster</i> , 1523 aa. [WO200171042-A2, 27-SEP-2001]	1..658 1..660	309/705 (43%) 412/705 (57%)	e-141
AAB43113	Human ORFX ORF2877 polypeptide sequence SEQ ID NO:5754 - <i>Homo sapiens</i> , 221 aa. [WO200058473-A2, 05-OCT-2000]	1171..1332 1..162	162/162 (100%) 162/162 (100%)	3e-87
AAB41768	Human ORFX ORF1532 polypeptide sequence SEQ ID NO:3064 - <i>Homo sapiens</i> , 128 aa. [WO200058473-A2, 05-OCT-2000]	683..801 2..121	115/120 (95%) 116/120 (95%)	6e-59

In a BLAST search of public sequence databases, the NOV37a protein was found to have homology to the proteins shown in the BLASTP data in Table 37D.

Table 37D. Public BLASTP Results for NOV37a				
Protein Accession Number	Protein/Organism/Length	NOV37a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9H3X8	DJ927M24.2 (KIAA1219) - <i>Homo sapiens</i> (Human), 1188 aa (fragment).	1..1169 1..1188	1143/1208 (94%) 1144/1208 (94%)	0.0
BAA86533	KIAA1219 protein - <i>Homo sapiens</i> (Human), 1112 aa (fragment).	651..1332 371..1053	674/683 (98%) 677/683 (98%)	0.0
CAD39096	Hypothetical protein - <i>Homo sapiens</i> (Human), 1333 aa (fragment).	651..1332 591..1274	674/684 (98%) 677/684 (98%)	0.0

Q9ULK1	KIAA1219 protein - <i>Homo sapiens</i> (Human), 532 aa (fragment).	860..1332 1..473	473/473 (100%) 473/473 (100%)	0.0
Q8WWC0	Hypothetical 47.6 kDa protein - <i>Homo sapiens</i> (Human), 423 aa (fragment).	970..1332 2..364	363/363 (100%) 363/363 (100%)	0.0

Pfam analysis predicts that the NOV37a protein contains the domains shown in Table 37E.

Table 37E. Domain Analysis of NOV37a			
Pfam Domain	NOV37a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 38.

The NOV38 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 38A.

Table 38A. NOV38 Sequence Analysis			
	SEQ ID NO: 227	3116 bp	
NOV38a, CG97012-01 DNA Sequence	<p>ATGCCCCCGCCCGGCCCCCGCCCGCCGCGCTGCGGGGCATCAGCCTGTTCTTGGCCCTG CTGCTGGGCAGCCCGCCCGCCCGCTGGAGCGGGACGCCCTGCCGAGGGCGACGCCAGC CCCCCTGGGCCCCCTACCTGTGCTGCCAGCGCGCCCCGAGCGGGGAGCCCCGGCAAGGAG CACCCCGAGGAGCGGGTGGTGACCGCCCCCCCCAGCAGCAGCCAGAGCGCCGAGGTGCTG GGCGAGCTGGTGTGGACGGCACC GCCCCCCAGCGCCCAACACGACATCCCCGCCCTGAGC CCCCTGCTGCCCCGAGGAGGCCCGGCCAAGCACGCCCTGCCCCCAAGAAGAAGCTGCCCC AGCCTGAAGCAGGTGAACAGCGCCCGGAAGCAGCTGCGGCCCAAGGCCACCGCCGCGCC ACCGTGACAGCGGGCCGGCAGCCAGCCCCGCCAGCCAGGGCCTGGACCTGCTGAGCAGCAGC ACCGAGAAGCCCGGCCCCCCCCGGCGACCCCGACCCCATCGTGGCCAGCGAGGAGGCCAGC GAGGTGCCCCCTGTGGCTGGACCGGAAGGAGAGCGCCGTGCCACCAACCCCGCACCCTTG CAAATCTCCCCCTTCACTTCGCAGCCCTATGTGGCCACACACTCCCCCAGAGGCCAGAA CCCGGGGAGCCTGGGCCTGACATGGCCAGGAGGCCCCCCAGGAGGACACCGCCCATG GCCCTGATGGACAAAGGTGAGAATGAGCTGACTGGGTGAGCCTCAGAGGAGAGCCAGGAG ACCACTACCTCCACCATATCACCACCAGGTATCACCACCGAGCAGGCACCACTCTC TGCACTGTGAGCTTCTCAATCCTGAGGGGTACATGACTCCAGCGACTACCACTGCTG CCCCCAACAACCTTTCTGGAGTGACATACAACGTGACAGTTACACTGGCTATGGGGTG GAGCTCCAGGTGAAGAGTGTGAACCTGTCCGATGGGGAAGTGTCTCCATCCGCGGGGTG GACGGCCCTACCTGACCGTCTGGCCAACAGACACTCCTGGTGGAGGGGAGGTAATC CGAAGCCCCACCAACACCATCTCCGTCTACTTCCGGACCTTCCAGGACGACGGCCTGGG ACCTTCCAGCTTCACTACCAGGCCTTCATGCTGAGCTGCAACTTTCCCCCGCGCCTGAC TCTGGGGATGTACGGTGATGGACCTGCACTCAGGTGGGGTGGCCCACTTCACTGCCAC CTGGGCTATGAGCTCCAGGGCGCTAAGATGCTGACATGCATCAATGCCTCAAGCCGCAC TGGAGCAGCCAGGAGCCCATCTGCTCAGCTCCTTGTGGAGGGGAGTGCACAATGCCACC ATCGCCCGCTCTCTCCCAAGTTACCTGAAACACCAATGGGAGGCAATCTGCATC TGGACGATTGAAGCTCCAGAGGGCCAGAAGCTGCACCTGCACTTTGAGAGGCTGTTGCTG CATGACAAGGACAGGATGACGGTTCACAGCGGGCAGACCAACAAGTCAGCTCTTCTCTAC GACTCCCTTCAAACCGAGAGTGTCCCTTTTGAGGGCCTGCTGAGCGAAGGCAACACCATC CGCATCGAGTTCACGTCCGACCAGGCCCGGGCGGCCTCCACCTTCAACATCCGATTGAA</p>		

	GCGTTTGAGAAAGGCCACTGCTATGAGCCCTACATCCAGAATGGGAACCTCACTACATCC GACCCGACCTATAACATTGGGACTATAGTGGAGTTCACCTGCGACCCCGGCCACTCCCTG GAGCAGGGCCCCGCCATCATCGAATGCATCAATGTGCGGGACCCATACTGGAATGACACA GAGCCCCTGTGCAGAGCCATGTGTGGTGGGGAGCTCTCTGCTGTGGCTGGGGTGGTATTG TCCCCAAACTGGCCCCGAGCCCTACGTGGAAGGTGAAGATTGTATCTGGAAGATCCACGTG GGAGAAGAGAAACGGATCTTCTTAGATATCCAGTTCCTGAATCTGAGCAACAGTGACATC TTGACCATCTACGATGGCGACGAGGTTCATGCCCCACATCTTGGGGCAGTACCTTGGGAAC AGTGGCCCCCAGAACTGTACTCCTCCACGCCAGACTTAACCATCCAGTTCATTTCGGAC CCTGCTGGCCTCATCTTTGAAAGGGCCAGGGATTATCATGAACTACATAGAGGTATCA AGGAATGACTCCTGCTCGGATTTACCCGAGATCCAGAATGGCTGGAAAACCACTTCTCA ACGGAGTTGGTGGCGGGGAGCCAGAATCACTACCAGTGTGACCCCGGCTATGACATCGT GGGAGTGACACCCTCACCTGCCAGTGGGACCTCAGCTGGAGCAGCGACCCCCCATTTTGT GAGAAAATTATGTACTGCACCGACCCCGAGAGGTGGATCACTCGACCCGCTTAATTTTCG GATCCTGTGCTGCTGGTGGGGACCACTCCAATACACCTGCAACCCCGGTTTTGTGCTT GAAGGGAGTTCTCTTCTGACCTGCTACAGCCGTGAAACAGGGAATCCCATCTGGACGTCT CGCCTGCCCCACTGCGTTTTCGAGGAGTCCCTGGCATGTGACAACCCAGGGCTGCCTGAA AATGGATACCAATCCTGTACAAGCGACTCTACCTGCCAGGAGAGTCCCTCACCTTCATG TGCTACGAAGGCTTTGAGCTCATGGGTGAAGTGACCATCCGCTGCATCCTGGGACAGCCA TCCCCTGGAACGGGCCCCTGCCCGTGTGTAAAGTAGCAGAAGCGGCAGCAGAGACGTCTG CTGGAAGGGGGGAACATGGCCCTGGCTATCTTCATCCGGTCCCTCATCTCCTTACTG CTGGAGGAGCCTACATTTACATCACAGATGTGCTACTATTCCAACCTCCGCTGCCT CTGATGTACTCCACCCCTACAGCCAGATCACCGTGGAAACCGAGTTTGACAACCCCAT TACGAGACAGGGGAAACCAGAGAGTATGAGGTTTCTATCTAAAGAGAGCTACACTTGAGA AGGGGACTTGTGAACTCAACCACAATCTCCTCGAGACATTATCCAGAGACCATGT		
	ORF Start: ATG at 1		ORF Stop: TAA at 3040
	SEQ ID NO: 228	1013 aa	MW at 110509.9kD
NOV38a, CG97012-01 Protein Sequence	MPAARPPAAGLRGISLFLALLLGSPAAALERDALPEGDASPLGPYLLPSGAPERGSPGKE HPEERVVTAPSSSSQSAEVLGELVLDGTAPSAHDI PALSPLLPEEARPKHALPPKKLP SLKQVNSARKQLRPKATSAAVTQVQAGSQPASQGLDLLSSSTEKPGPPGDPDPVASEEAS EVPLWLDKESAVPTTPAPLQISPFTSQPYVAHTLPORPEPGEPPGDMQEAPEQEDTSPM ALMDKGENELTGSASEESQETTTSTIIITTVITTEQAPALCSVSFSNPEGYIDSDYPLL PLNNFLECTYNVTYVYGYGVELQVKS VNLSDGELLSIRGVDGPTLTVLANQTLLEVGQVI RSPNTTISVYFRTFQDDGLGTFQLHYQAFMLSCNFRRPDSGDVTVMDLHSGGVAHFHCH LGYELQGAKMLTCINASKPHWSSQEPICSA PCGGAVHNATIGRVLSPSYPENTNGSQFCI WTIEAPEGQKLHLHFERLLLHDKDRMTVHSGQTNKSALLYDSLQTESVPFEGLLSEGNTI RIEFTSDQARAASTFNIRFEAFKHCYEPYIQNGNFTTSDPTYNIGTIVEFTCDPGHSL EQGPAIECINVRDPYWNDEPLCRAMCGGELS AVAGVVLSPNWPEPYVEGEDCIWKIHV GEEKRIFLDIQFLNLSNSDILTIYDGDEVMPHILGQYLGNSGPQKLYSSTPDLTIQFHSD PAGLIFGKGQGFIMNYIEVSRNDSCLPEIQNGWKTTSHTELVRGARITYQCDPGYDIV GSDTLTCQWDLSSWSDPPFCEKIMYCTDPGEVDHSTR LISDPVLLVGTIIQYTCNPGFVL EGSSLLTCYSRETGTPIWTSRLPHCVSEESLACDNPLGPENGYQILYKRLYLPGESLTFM CYEGFELMGEVTIRCILGQPSHWNGPLPVCKVAEAAAETSLEGGNMALAI FIPVLIISLL LGGAYIYITRCRYYSNLRLPLMYSHPYSQITVETEFDNPIYETGETREYEVSI		
	SEQ ID NO: 229	2420 bp	
NOV38b, CG97012-02 DNA Sequence	CCTGGGCTGACATGGCCCAGGAGGCCCCAGGAGGACACCAGCCCCATGGCCCTGATG GACAAAGGTGAGAATGAGCTGACTGGGTGAGCCTCAGAGGAGAGCCAGGAGACCCTACC TCCACCATATCACCACCACGGTCATCACCACCAGCAGGCACCAGCTCTCTGCAGTGTG AGCTTCTCCAATCCTGAGGGGTACATTGACTCCAGCGACTACCACTGGTGGCCCTCAAC AACTTCTCGAGTGACATACAACGTGACAGTCTACACTGGCTATGGGGTGGAGCTCCAG GTGAAGAGTGTGAACCTGTCCGATGGGGAAGTGTCTCCATCCGCGGGGTGGACGGCCCT ACCCTGACCGTCTCTGGCCAACAGACACTCCTGGTGGAGGGGCAGGTAATCCGAAGCCCC ACCACACCATCTCCGTCTACTTCCGACCTTCCAGGACGACGGCCTTGGGACCTTCCAG CTTCACTACCAGGCCTTCATGCTGAGCTGCAACTTTCCCCGCCGGCCTGACTCTGGGGAT GTCACGGTGATGGACCTGCACTCAGGTGGGGTGGCCCACTTCACTGCCACCTGGGCTAT GAGCTCCAGGGCGCTAAGATGCTGACATGCATCAATGCCTCCAAGCCGCACTGGAGCAGC CAGGAGCCCCTCTGCTCAGCTCCTTGTGGAGGGGCAGTGCACAATGCCACCATCGGCCGC GTCCTCTCCCAAGTTACCTGAAACACAAATGGGAGCCAATTCTGCATCTGGACGATT GAAGCTCCAGAGGGCCAGAAGCTGCACCTGCACCTTGAGAGGCTGTTGCTGCATGACAAG		

	GACAGGATGACGGTTCACAGCGGGCAGACCAACAAGTCAGCTCTTCTCTACGACTCCCTT CAAACCGAGAGTGTCCCTTTTGAGGGCCTGCTGAGCGAAGGCAACACCATCCGCATCGAG TTCACGTCCGACCAGGCCCGGGCGGCCTCCACCTTCAACATCCGATTTGAAGCGTTTGAG AAAGGCCACTGCTATGAGCCCTACATCCAGAATGGGAACCTTCACTACATCCGACCCGACC TATAACATTGGGACTATAGTGGAGTTCACCTGCGACCCCGGCCTCCCTGGAGCAGGGC CCGGCCATCATCGAATGCATCAATGTGCGGGACCCATACTGGAATGACACAGAGCCCCTG TGCAGAGCCATGTGTGGTGGGAGCTCTCTGCTGTGGCTGGGGTGGTATTGTCCCCAAC TGGCCCGAGCCCTACGTGGAAGGTGAAGATGTATCTGGAAGATCCACGTGGGAGAAGAG AAACGGATCTTCTTAGATATCCAGTTCCTGAATCTGAGCAACAGTGACATCTTGACCATC TACGATGGCGACGAGGTATGCCCCACATCTTGGGGCAGTACCTTGGGAACAGTGGCCCC CAGAAACTGTACTCCTCCACGCCAGACTTAACCATCCAGTTCATTCCGACCCCTGCTGGC CTCATCTTTGGAAGGGCCAGGGATTTATCATGAACTACATAGAGGTATCAAGGAATGAC TCCTGCTCGGATTTACCCGAGATCCAGAATGGCTGGAAGAACCTTCTCACACGGAGTTG GTGCGGGGAGCCAGAATCACCTACCAGTGTGACCCCGCTATGACATCGTGGGGAGTGAC ACCTCACCTGCCAGTGGGACCTCAGCTGGAGCAGCAGCCCCCATTTTGTGAGAAAATT ATGTACTGCACCGACCCCGAGAGGTGGATCACTCGACCCGCTTAATTTCCGATCCTGTG CTGCTGGTGGGACCACCATCCAATACACCTGCAACCCCGGTTTTGTGCTTGAAGGGAGT TCTCTTCTGACCTGCTACAGCCGTGAAACAGGGACTCCCATCTGGACGTCTCGCCTGCCC CACTGCGTTTTCCGAGGAGTCCCTGGCATGTGACAACCCAGGGCTGCCTGAAAATGGATAC CAAATCCTGTACAAGCGACTCTACCTGCCAGGAGAGTCCCTCACCTTCATGTGCTACGAA GGCTTTGAGCTCATGGGTGAAGTGACCATCCGCTGCATCCTGGGACAGCCATCCCACTGG AACGGGCCCCCTGCCCGTGTGTAAAGTTAATCAAGACAGTTTTTGAACATGCTTTAGAAGTA GCAGAAGCGGCAGCAGAGACGTGCTGGAAGGGGGGAACATGGCCCTGGCTATCTTCATC CCGGTCTCATCATCTCCTTACTGCTGGGAGGAGCCTACATTACATCAAGATGTGCG TACTATTCCAACCTCCGCCTGCCTCTGATGTACTCCACCCCTACAGCCAGATCACCGTG GAAACCGAGTTTGACAACCCCATTTACGAGACAGGGGAAACCAGAGAGTATGAGGTTTCT ATCTAAAGAGAGCTACACTT		
	ORF Start: ATG at 13		ORF Stop: TAA at 2404
	SEQ ID NO: 230	797 aa	MW at 88285.1kD
NOV38b, CG97012-02 Protein Sequence	MAQEAPQEDTSPMALMDKGENELTGSASEBSQETTTSTIIITTVITTEQAPALCSVSFSN PEGYIDSSDYPLPLNPFLECTYNVTVTGYGVQLQKSVNLSGDELLSIRGVDGPTLT LANQTLLEVGQVIRSPNTISVYFRTFQDDGLGTFQLHYQAFMLSCNFPFRPDSDGVTVM DLHSGGVAHFHCHLGYELQGAQMLTLCINASKPHWSSQEPICSAFCGGAVHNATIGRVLS SYPENTNGSQFCIWTIEAPEGQKLHLHFERLLLHDKDRMTVHSGQTNKSALLYDSLQTES VPFEGLLSEGNTRIEFTSDQARAASTFNIRFEAFKGHCEPYIQNGNFTTSDPTYNIG TIVEFTCDPGHSLQGPALIECINVRDPYWNDEPLCRAMCGGELSAGAVVLSNPWPEP YVEGEDCIWKIHWGEEKRIFLDIQFLNLSNSDILTIYDGDEVMPHILGQYLGNSGPQKLY SSTPDLTIQFHSDFAGLIFGKGQGFIMNYIEVSRNDSCLDPEIQNGWKTTSHTELVRGA RITYQCDPGYDIVGSDTLTCQWDLSSWSDPPFCEKIMYCTDPGEVDHSTRILSDPVLVVG TTIYTCNPGFVLEGGSSLLTCYSRETGTPIWTSRLPHCVSEESLACDNPLGPENGYQILY KRLYLPGESLTFMCEYGFELMGEVTIRICILGQPSHWNGPLPVCKVNQDSFEHALEVAEAA AETSLEGGNMALAI FIPVLIISLLGGAYIYITRCRYSNLRLPLMYSHPSYQITVETEF DNPIYETGETREYEVS I		
	SEQ ID NO: 231	1434 bp	
NOV38c, CG97012-03 DNA Sequence	AGATCTTGCAACTTTCCCCCGCGCCTGACTCTGGGGATGTACGGTGTGACCTGCAC TCAGGTGGGGTGGCCCACTTCACTGCCACCTGGGCTATGAGCTCCAGGGCGCTAAGATG CTGACATGCATCAATGCCTCCAAGCCGCACTGGAGCAGCCAGGAGCCCATCTGCTCAGCT CCTTGTGGAGGGGCACTGCACAATGCCACCATCGGCCCGCTCCTCTCCCAAGTTACCTT GAAAACACCAATGGGAGCCAATTCTGCATCTGGACGATTGAAGCTCCAGAGGGCCAGAAG CTGCACCTGCACTTTGAGAGGCTGTTGCTGCATGACAAGGACAGGATGACGGTTACAGC GGCAGACCAACAAGTCAGCTCTTCTCTACGACTCCCTTCAAACCGAGAGTGTCCCTTTT GAGGGCCTGCTGAGCGAAGGCAACACCATCCGCATCGAGTTCACGTCCGACCAGGCCCCG GCGGCCTCCACCTTCAACATCCGATTGAAGCGTTTGAGAAAGGCCACTGCTATGAGCCC TACATCCAGAATGGGAACCTACATACATCCGACCCGACCTATAACATATGGGACTATAGT GAGTTACCTGCGACCCCGGCCACTCCTGGAGCAGGGCCCGGCCATCATCGAATGCATC AATGTGCGGGACCCATACTGGAATGACACAGAGCCCTGTGCAGAGCCATGTGTGGTGGG GAGCTCTCTGCTGTGGCTGGGGTGGTATTGTCCCAAAGTGGCCCGAGCCCTACGTGGAA GGTGAAGATTGTATCTGGAAGATCCAGTGGGAGAAGAGAAACGGATCTTCTAGATATC		

	CAGTTCCTGAATCTGAGCAACAGTGACATCTTGACCATCTACGATGGCGACGAGGTCATG CCCCACATCTTGGGGCAGTACCTTGGGAACAGTGGCCCCAGAACTGTACTCCTCCAG CCAGACTTAACCATCCAGTTCATTCCGACCCTGCTGGCCTCATCTTTGGAAGGGCCAG GGATTATCATGAACACATAGAGGTATCAAGGAATGACTCCTGCTCGGATTTACCCGAG ATCCAGAATGGCTGGAAAACCACTTCTCACACGGAGTTGGTGGGGGAGCCAGAATCACC TACCAGTGTGACCCCGGCTATGACATCGTGGGGAGTGACACCTCACCTGCCAGTGGGAC CTCAGCTGGAGCAGCGACCCCCCATTTTGTGAGAAAACGGAGGAGTCCCTGGCATGTGAC AACCAGGGGCTGCTGAAAATGGATACCAAATCCTGTACAAGCGACTCTACCTGCCAGGA GAGTCCCTCACCTTCATGTGCTACGAAGGCTTTGAGCTCATGGGTGAAGTGACCATCCGC TGCATCCTGGGACAGCCATCCCCTGGAACGGGGCCCTGCCCGTGTGTGTCGAC		
	ORF Start: at 7		ORF Stop: at 1429
	SEQ ID NO: 232	474 aa	MW at 52744.6kD
NOV38c, CG97012-03 Protein Sequence	CNFPRRPDSGDVTVMDLHSGGVAHFHCHLGYELQGAKMLTCINASKPHWSSQEPICSA PGAVHNATIGRVLSPSYPTNGSQFCIWTIEAPEGQKLHLHFERLLLDHDKDRMTVHSGQ TNKSALLYDSLQTESVPFEGLLSEGTIRIEFTSDQARAASTFNIRFEAFEGKHCEPYI QNGNFTTSDPTYNIGTIVEFTCDPGHSLEQGPALIECINVRDPYWNDEPLCRAMCGGEL SAVAGVVLSPNWPEPYVEGEDCIWKIHVGEEKRIFLDIQFLNLSNSDILTIYDGDEVMPH ILGQYLGNSGPQKLYSSTPDLTIQFHSDPAGLIFGKGQGFIMNYIEVSRNDSCLDPEIQ NGWKTTSHTELVRGARITYQCDPQYDIVGSDTLTCQWDLWSWSDPPFCKTEESLACDNP GLPENGQILYKRLYLPGESLTFCYEGFELMGEVTIRCILGQPSHWNGPLPVC		
	SEQ ID NO: 233	3116 bp	
NOV38d, CG97012-01 DNA Sequence	ATGCCCGCCGCGCCGCCCCCGCCGCGCCTGCGGGGCATCAGCCTGTTCTGGCCCTG CTGCTGGGCAGCCCCGCGCGCCCTGGAGCGGGACGCCCTGCCGAGGGCGACGCCAGC CCCCGAGGAGCGGGTGGTACCGCCCCCCCCAGCAGCAGCCAGAGCGCCGAGGTGCTG GGCGAGCTGGTGTGGACGGCACCAGCCCCAGCGCCACACGACATCCCCGCCCTGAGC CCCCGCTGCTGCCGAGGAGGCCCGGCCAAGCAGCCCTGCCCCCCAAGAAGAAGCTGCC AGCCTGAAGCAGGTGAACAGCGCCCCGGAAGCAGCTGCGGCCCAAGGCCACCGCCGCC ACCGTGACGCGGGCCGGCAGCCAGCCCCGCCAGCCAGGGCCTGGACCTGCTGAGCAGCAGC ACCGAGAAGCCCCGCCCCCGCGGACCCCGACCCCATCGTGGCCAGCAGGAGGCCAGC GAGGTGCCCCCTGTGGCTGGACCGGAAGGAGAGCGCGTGGCCACCACCCCCGACCCCTG CAAATCTCCCCCTTCACTTCGACGCCCTATGTGGCCACACACTCCCCAGAGGCCAGAA CCCCGGGAGCCTGGGCCTGACATGGCCAGGAGGCCCCCCAGGAGGACACAGCCCCATG GCCCTGATGGACAAAGGTGAGAATGAGCTGACTGGGTGAGCCTCAGAGGAGAGCCAGGAG ACCACTACCTCCACCATTATCACCACCAGGTATCACCACCGAGCAGGACACAGCTCTC TGCACTGTGAGCTTCTCCAATCCTGAGGGGTACATTGACTCCAGCGACTACCACTGCTG CCCCCAACAATTTCTGGAGTGACATACAACGTGACAGTCTACACTGGCTATGGGGTG GAGCTCCAGGTGAAGAGTGTGAACCTGTCCGATGGGGAAGTGTCTCTCATCCGCGGGGTG GACGGCCCTACCTGACCGTCTTGCCCAACAGACACTCCTGGTGGAGGGGAGGTAATC CGAAGCCCCACCAACACCATCTCCGTCTACTTCCGACCTTCCAGGACGAGGCCCTGGG ACCTTCCAGCTTCACTACCGGCCTTCATGCTGAGCTGCAACTTTCCCCGCGGCCTGAC TCTGGGATGTACGGTGATGGACCTGCACTCAGGTGGGGTGGCCACTTCACTGCCAC CTGGGCTATGAGCTCCAGGGCGCTAAGATGCTGACATGCATCAATGCCTCCAAGCCGAC TGGAGCAGCCAGGAGCCCATCTGCTCAGCTCCTTGTGGAGGGGAGTGACAATGCCACC ATCGGCCGCGTCTCTCCCCAAGTTACCCTGAAAACACCAATGGGAGCCAATTCTGCATC TGGACGATTGAAGCTCCAGAGGGCCAGAAGCTGCACCTGCACCTTTGAGAGGCTGTTGCTG CATGACAAGGACAGGATGACGGTTCACAGCGGGCAGACCAACAAGTCAGCTCTTCTCTAC GACTCCCTTCAAACCGAGAGTGTCCCTTTTGGAGGGCCTGCTGAGCGAAGGCAACACCATC CGCATCGAGTTCACGTCCGACCAGGCCCGGGCGGCCTCCACCTTCAACATCCGATTTGAA GCGTTTGAGAAAGGCCACTGCTATGAGCCCTACATCCAGAATGGGAAGTCACTACATCC GACCCGACCTATAACATTGGGACTATAGTGGAGTTCACCTGCGACCCCGGCCACTCCTG GAGCAGGGCCCGCCATCATCGAATGCATCAATGTGCGGGACCCATAGGAATGACACA GAGCCCTGTGCGAGGACCATGTGTGGTGGGGAGCTCTCTGCTGTGGCTGGGGTGGTATTG TCCCCAAACTGGCCCGAGCCCTACGTGGAAGGTGAAGATTGTATCTGGAAGATCCACGTG GGAGAAGAGAAACGGATCTTCTTAGATATCCAGTTCCTGAATCTGAGCAACAGTGACATC TTGACCATCTACGATGGCGACGAGGTATGCCCCACATCTTGGGGCAGTACCTTGGGAAC AGTGGCCCCCAGAAACTGTACTCCTCCAGCCAGACTTAACCATCCAGTTCATTCCGAC CCTGCTGGCCTCATCTTTGAAAGGGCCAGGGATTATCATGAACACATAGAGGTATCA		

	AGGAATGACTCTGCTCGGATTTACCCGAGATCCAGAATGGCTGGAAAACCACTTCTCAC ACGGAGTTGGTGCAGGGGAGCCAGAATCACCTACCAGTGTGACCCCGGCTATGACATCGTG GGGAGTGACACCCTCACCTGCCAGTGGGACCTCAGCTGGAGCAGCGACCCCCATTGTTGT GAGAAAATTATGTAAGTGCACCGACCCCGGAGAGGTGGATCACTCGACCCGCTTAATTTCCG GATCCTGTGCTGCTGGTGGGGACCACCATCCAATACACCTGCAACCCGGTTTTGTGCTT GAAGGGAGTTCTCTTCTGACCTGCTACAGCCGTGAAACAGGAGACTCCCATCTGGACGTCT CGCCTGCCCCACTGCGTTTTCGGAGGAGTCCCTGGCATGTGACAACCCAGGGCTGCCTGAA AATGGATACCAAATCCTGTACAAGCGACTCTACCTGCCAGGAGAGTCCCTCACCTTCATG TGCTACGAAGGCTTTGAGCTCATGGGTGAAGTGACCATCCGCTGCATCTCTGGGACAGCCA TCCCACTGGAACGGGCCCCCTGCCCGTGTGTAAGTAGCAGAAGCGGCAGCAGAGACGTCC CTGGAAGGGGGGAACATGGCCCTGGCTATCTTCATCCCGGTCCTCATCATCTCCTTACTG CTGGGAGGAGCCTACATTTACATCACAAGATGTCGCTACTATTCCAACCTCCGCTGCCT CTGATGTACTCCACCCCTACAGCCAGATCACCGTGGAAACCGAGTTTGACAACCCCAT TACGAGACAGGGGAAACAGAGAGTATGAGGTTTCTATCTAAGAGAGCTACACTTGAGA AGGGGACTTGTGAATCAACCACAATCTCCTCGAGACATTATCCAGAGACCATGT		
	ORF Start: ATG at 1		ORF Stop: TAA at 3040
	SEQ ID NO: 234	1013 aa	MW at 110509.9kD
NOV38d, CG97012-01 Protein Sequence	MPAARPPAAGLRGISLFLALLLGSPAAALERDALPEGDASPLGPYLLPSGAPERGSPGKE HPEERVVTAPPSSQSAEVLGELVLDGTAPSAHHDIPALSPLLPEEARPKHALPPKKLP SLKQVNSARKQLRPKATSAAATVQRAGSQPASQGLDLLSSSTEKPGPPGDPDIVASEEAS EVPLWLDRKESAVPTTPAPLQISPFTSQPYVAHTLPQRPEPEGEPGDMAQEAPQEDTSPM ALMDKGENELTGSASEESQETTTSTIIITTVITTEQAPALCSVSFSNPEGYIDSSDYPLL PLNMFLECTYNTVYTYGYVGLQVKS VNLS DGE LLSIRGVDGPTLTVLANQTL LVEGQVI RSPTNTISVYFRTFQDDGLGTFQLHYQAFMLSCNFPRRPDSGDVTVM DLHSGGVAFHCH LGYELQGA KMLTCINASKPHWSSQEPICSA PCGGA VHNATIGRVLSPSY PENTNGSQFCI WTIEAPEGQKLHLHFERLLLDKDRMTVHSGQTNKSALLYDSLQTESVPFEGLLSEGTI RIEFTSDQARAASFTNIRFEAFEGHCHYEPYIQNGNFTTSDPTYNIGTIVEFTCDPGHSL EQGPALIECINVRDPYWNDEPLCRAMCGGELS AVAGVVLSPNWPEPYVEGEDCIWKI HV GEEKRIFLDIQFLNLSNSDILTIYDGDVMPHILGQYLGNSGPQKLYSSTPDLTIQFHSD PAGLIFGKGQGFIMNYIEVSRNDSCLPEIQNGWKTTSHTELVRGARITYQCDPGYDIV GSDTLTCQWDLWSDDPPFCEKIMYCTDPGEVDHSTRLLISDPVLLVGTTIQTCTNPGFVL EGSSLLTCYSRETGTPIWTSRLPHCVSEESLACDNPLPENGYQILYKRLYLPGESLTFM CYEGFELMGEVTIRCILGQPSHWNGPLPVCKVAEAAAETSLEGGNMALAFIPVLIISLL LGGAYIYITRCRYNSNRLPLMYSHPYSQITVETFDNPIYETGETREYEVSI		
	SEQ ID NO: 235		867 bp
NOV38e, 210120300 DNA Sequence	AGATCTTGTGGAGGGGCGAGTGCACAATGCCACCATCGGCCGCGTCCTCTCCCAAGTTAC CCTGAAAACACCAATGGGAGCCAATTCTGCATCTGGACGATTGAAGCTCCAGAGGGCCAG AAGCTGCACCTGCACCTTGAGAGGCTGTGCTGCATGACAAGGACAGGATGACGGTTCAC AGCGGGCAGACCAACAAGTCACTCTTCTCTACGACTCCCTTCAAACCGAGAGTGTCCCT TTTGAGGGCCTGCTGAGCGAAGGCAACACCATCCGCATCGAGTTCACGTCCGACAGGCC CGGGCGGCCTCCACCTTCAACATCCGATTTGAAGCGTTTGAGAAAGGCCACTGCTATGAG CCCTACATCCAGAATGGGAACCTCACTACATCCGACCCGACCTATAACATTGGGACTATA GTGGAGTTACCTGCGACCCCGGCCACTCCCTGGAGCAGGGCCCGGCATCATCGAATGC ATCAATGTGCGGGACCCATACTGGAATGACACAGAGCCCTGTGCAGAGCCATGTGTGGT GGGGAGCTCTCTGCTGTGGCTGGGGTGGTATTGTCCCCAAACTGGCCCGAGCCCTACGTG GAAGGTGAAGATTGTATCTGGAAGATCCACGTGGGAGAAGAGAAACGGATCTTCTTAGAT ATCCAGTTCCTGAATCTGAGCAACAGTGACATCTTGACCATCTACGATGGCGACGAGGTC ATGCCCCACATCTTGGGGCAGTACCTTGGGAACAGTGGCCCCAGAACTGTACTCCTCC ACGCCAGACTTAACCATCCAGTTCATTCCGACCCTGCTGGCCTCATCTTTGGAAGGGC CAGGGATTTATCATGAACCTACGTCGAC		
	ORF Start: at 1		ORF Stop: end of sequence
	SEQ ID NO: 236	289 aa	MW at 32172.6kD
NOV38e, 210120300 Protein Sequence	RSCGAVHNATIGRVLSPSY PENTNGSQFCIWTIEAPEGQKLHLHFERLLLDKDRMTVH SGQTNKSALLYDSLQTESVPFEGLLSEGTIRIEFTSDQARAASFTNIRFEAFEGHCHYE PYIQNGNFTTSDPTYNIGTIVEFTCDPGHSL EQGPALIECINVRDPYWNDEPLCRAMCG GELS AVAGVVLSPNWPEPYVEGEDCIWKI HVGEEKRIFLDIQFLNLSNSDILTIYDGDV MPHILGQYLGNSGPQKLYSSTPDLTIQFHSDPAGLIFGKGQGFIMNYVD		

	SEQ ID NO: 237	867 bp
NOV38f, 210120376 DNA Sequence	AGATCTTGTGGAGGGGCAGTGCACAATGCCACCATCGGCCGCGTCTCTCCCCAAGTTAC CCTGAAACACAAATGGGAGCCAATTCTGCATCTGGACGATTGAAGCTCCAGAGGGCCAG AAGCTGCACCTGCACCTTTGAGAGGCTGTTGCTGCATGACAAGGACAGGATGACGGTTCAC AGCGGGCAGACCAACAAGTCAGCTCTTCTCTACGACTCCCTTCAAACCGAGAGTGTCCCT TTTGAGGGCCTGCTGAGCGAAGGCAACACCATCCGCATCGAGTTCACGTCCGACCAGGCC CGGGCGGCTCCACCTTCAACATCCGATTGAAGCGTTTGAGAAAGGCCACTGCTATGAG CCCTACATCCAGAATGGGAACCTTCACTACATCCGACCCGACCTATAACATTGGGACTATA GTGGAGTTCACCTGCGACCCCGGCCACTCCCTGGAGCAGGGCCCGGCCATCATCGAATGC ATCAATGTGCGGGACCCATACTGGAATGACACAGAGCCCCTGTGCAGAGCCATGTGTGGT GGGGAGCTCTCTGCTGTGGCTGGGGTGGTATTGTCCCCAACTGGCCCGAGCCCTACGTG GAAGGTGAAGATTGTATCTGGAAGATCCACGTGGGAGAAGAGAAACCGATCTTCTAGAT ATCCAGTTCCTGAATCTGAGCAACAGTGACATCTTGACCATCTACGATGGCGACGAGGTC ATGCCCCACATCTTGGGGCAGTACCTTGGGAACAGTGGCCCCCAGAACTGTACTCCTCC ACGCCAGACTTAACCATCCAGTTCATTGCGACCCTGCTGGCCTCATCTTTGAAAGGGC CAGGGATTATCATGAACCTACGTCGAC	
	ORF Start: at 1	ORF Stop: end of sequence
	SEQ ID NO: 238	289 aa MW at 32172.6kD
NOV38f, 210120376 Protein Sequence	RSCGGAHVHNTIGRVLSPSPYENTNGSQFCIWTIEAPEGQKLHLHFERLLLHDKDRMTVH SGQTNKSALLYDSLQTESVPFEGLLSEGNIRIEFTSDQARAASFNIRFEAFKKGHCYE PYIQNGNFTTSDPTYNIGTIVEFTCDPGHSLEQGPALIECINVRDPYWNDEPLCRAMCG GELSAVAGVVLSPNWPEPYVEGEDCIWKIHVGEEKRIFLDIQFLNLSNSDILTIYDGDEV MPHILGQYLGNSGPQKLYSSPDLTIQFHSDPAGLIFGKGQGFIMNYVD	
	SEQ ID NO: 239	867 bp
NOV38g, 210120463 DNA Sequence	AGATCTTGTGGAGGGGCAGTGCACAATGCCACCATCGGCCGCGTCTCTCCCCAAGTTAC CCTGAAACACCAATGGGAGCCAATTCTGCATCTGGACGATTGAAGCTCCAGAGGGCCGG AAGCTGCACCTGCACCTTTGAGAGGCTGTTGCTGCATGACAAGGACAGGATGACGGTTCAC AGCGGGCAGACCAACAAGTCAGCTCTTCTCTACGACTCCCTTCAAACCGAGAGTGTCCCT TTTGAGGGCCTGCTGAGCGAAGGCAACACCATCCGCATCGAGTTCACGTCCGACCAGGCC CGGGCGGCTCCACCTTCAACATCCGATTGAAGCGTTTGAGAAAGGCCACTGCTATGAG CCCTACATCCAGAATGGGAACCTTCACTACATCCGACCCGACCTATAACATTGGGACTATA GTGGAGTTCACCTGCGACCCCGGCCACTCCCTGGAGCAGGGCCCGGCCATCATCGAATGC ATCAATGTGCGGGACCCATACTGGAATGACACAGAGCCCCTGTGCAGAGCCATGTGTGGT GGGGAGCTCTCTGCTGTGGCTGGGGTGGTATTGTCCCCAACTGGCCCGAGCCCTACGTG GAAGGTGAAGATTGTATCTGGAAGATCCACGTGGGAGAAGAGAAACCGATCTTCTAGAT ATCCAGTTCCTGAATCTGAGCAACAGTGACATCTTGACCATCTACGATGGCGACGAGGTC ATGCCCCACATCTTGGGGCAGTACCTTGGGAACAGTGGCCCCCAGAACTGTACTCCTCC ACGCCAGACTTAACCATCCAGTTCATTGCGACCCTGCTGGCCTCATCTTTGAAAGGGC CAGGGATTATCATGAACCTACGTCGAC	
	ORF Start: at 1	ORF Stop: end of sequence
	SEQ ID NO: 240	289 aa MW at 32200.7kD
NOV38g, 210120463 Protein Sequence	RSCGGAHVHNTIGRVLSPSPYENTNGSQFCIWTIEAPEGRKLHLHFERLLLHDKDRMTVH SGQTNKSALLYDSLQTESVPFEGLLSEGNIRIEFTSDQARAASFNIRFEAFKKGHCYE PYIQNGNFTTSDPTYNIGTIVEFTCDPGHSLEQGPALIECINVRDPYWNDEPLCRAMCG GELSAVAGVVLSPNWPEPYVEGEDCIWKIHVGEEKRIFLDIQFLNLSNSDILTIYDGDEV MPHILGQYLGNSGPQKLYSSPDLTIQFHSDPAGLIFGKGQGFIMNYVD	
	SEQ ID NO: 241	1434 bp
NOV38h, 210120269 DNA Sequence	AGATCTTGCAACTTTCCCGCGCGCCTGACTCTGGGGATGTCACGGTGATGGACCTGCAC TCAGGTGGGGTGGCCCACTTTCACTGCCACCTGGGCTATGAGCTCCAGGGCGCTAAGATG CTGACATGCATCAATGCCTCAAGCCGCACTGGAGCAGCCAGGAGCCCATCTGCTCAGCT CCTTGTGGAGGGGCAGTGCACAATGCCACCATCGGCCGCGTCTCTCCCCAAGTTACCTT GAAAACACCAATGGGAGCCAATTCTGCATCTGGACGATTGAAGCTCCAGAGGGCCAGAAG CTGCACCTGCACCTTGGAGAGGCTGTTGCTGCATGACAAGGACAGGATACGGTTCACAGC GGGCAGACCAACAAGTCAGCTCTTCTCTACGACTCCCTTCAAACCGAGAGTGTCCCTTTT GAGGGCTGCTGAGCGAAGGCAACACCATCCGCATCGAGTTCACGTCCGACCAGGCCCCGG GCGGCTCCACCTTCAACATCCGATTGAAGCGTTTGAGAAAGGCCACTGCTATGAGCCC	

	TACATCCAGAATGGGAACCTTCACTACATCCGACCCGACCTATAACATTGGGACTATAGTG GAGTTCACCTGCGACCCCGGCCACTCCCTGGAGCAGGGCCCGGCATCATCGAATGCATC AATGTGCGGGACCCATACTGGAATGACACAGAGCCCTGTGCAGAGCCATGTGTGGTGGG GAGCTCTCTGCTGTGGCTGGGGTGGTATTGTCCCCAACTGGCCCGAGCCCTACGTGGAA GGTGAAGATTGTATCTGGAAGATCCACGTGGGAGAAGAGAAACGGATCTTCTTAGATATC CAGTTCCTGAATCTGAGCAACAGTGACATCTTGACCATCTACGATGGCGACGAGGTTCATG CCCCACATCTTGGGGCAGTACCTTGGGAAACAGTGGCCCCCAGAACTGTACTCCTCCACG CCAGACTTAACCATCCAGTTCATTCCGACCCCTGCTGGCCTCATCTTGGAAAGGGCCAG GGATTTATCATGAACACATAGAGGTATCAAGGAATGACTCCTGCTCGGATTTACCCGAG ATCCAGAATGGCTGGAAAACCACTTCTCACACGGAGTGGTGCGGGGAGCCAGAATCACC TACCAGTGTGACCCCGGCTATGACATCGTGGGGAGTGACACCCTCACCTGCCAGTGGGAC CTCAGCTGGAGCAGCGACCCCCATTTTGTGAGAAAACGGAGGAGTCCCTGGCATGTGAC AACCCAGGGCTGCCTGAAAATGGATACCAAATCCTGTACAAGCGACTCTACCTGCCAGGA GAGTCCCTCACCTTCATGTGCTACGAAGGCTTTGAGCTCATGGGTGAAGTGACCATCCGC TGCATCCTGGGACAGCCATCCCACTGGAACGGGGCCCTGCCCGTGTGTGTCGAC		
	ORF Start: at 1		ORF Stop: end of sequence
	SEQ ID NO: 242	478 aa	MW at 53202.0kD
NOV38h, 210120269 Protein Sequence	RSCNFRPRPDSGDVTVMDLHSGGVAHFHCHLGYELQGAKMLTCINASKPHWSSQEPICSA PCGGAVHNATIGRVLSPSPYENTNGSQFCIWTIEAPEGQKLHLHFERLLLHDKDRMTVHS GQTNKSALLYDSLQTESVPFEGLLSEGNTIRIEFTSDQARAASFTNIRFEAFKGHCYEP YIQNGNFTTSDPTYNIGTIVEFTCDPGHSLEQGPALIECINVRDPYWNDEPLCRAMCGG ELSAVAGVVLSPNWPEPYVEGEDCIWKIHVGEEKRIFLDIQFLNLSNSDILTIYDGDEV PHILGQYLGNSGPQKLYSSTPDLTIQFHSDPAGLIFGKGQGFIMNYIEVSRNDSCLDPE IQNGWKTTSHTELVRGARITYQCDPGYDIVGSDTLTCQWDLSSWSDPPFCKTEESLACD NPGLPENGYQILYKRLYLPGESLTFMCEYEGFELMGEVTIRCILGQPSHWNGPLPVCVD		
	SEQ ID NO: 243	867 bp	
NOV38i, CG97012-04 DNA Sequence	AGATCTTGTGGAGGGGAGTGCACAATGCCACCATCGGCCGCTCCTCTCCCCAAGTTAC CCTGAAAACACCAATGGGAGCCAATTCTGCATCTGGACGATTGAAGCTCCAGAGGGCCAG AAGCTGCACCTGCACCTTTGAGAGGCTGTTGCTGCATGACAAGGACAGGATGACGGTTCAC AGCGGGCAGACCAACAAGTCAGCTCTTCTCTACGACTCCCTTCAAACCGAGAGTGTCCCT TTTGAGGGCCTGCTGAGCGAAGGCAACACCATCCGCATCGAGTTCACGTCCGACCAAGGCC CGGGCGGCCTCCACCTTCAACATCCGATTTGAAGCGTTTGAGAAAGGCCACTGCTATGAC CCCTACATCCAGAATGGGAACCTTCACTACATCCGACCCGACCTATAACATTTGGGACTATA GTGGAGTTCACCTGCGACCCCGGCCACTCCCTGGAGCAGGGCCCGGCATCATCGAATGC ATCAATGTGCGGGACCCATACTGGAATGACACAGAGCCCTGTGCAGAGCCATGTGTGGT GGGGAGCTCTCTGCTGTGGCTGGGGTGGTATTGTCCCCAACTGGCCCCGAGCCCTACGTG GAAGGTGAAGATTGTATCTGGAAGATCCACGTGGGAGAAGAGAAACGGATCTTCTTAGAT ATCCAGTTCCTGAATCTGAGCAACAGTGACATCTTGACCATCTACGATGGCAGCAGGTC ATGCCCCACATCTTGGGGCAGTACCTTGGGAACAGTGGCCCCCAGAACTGTACTCCTCC ACGCCAGACTTAACCATCCAGTTCATTCCGACCCCTGCTGGCCTCATCTTGGAAAGGGC CAGGGATTTATCATGAACACGTCGAC		
	ORF Start: at 7		ORF Stop: at 862
	SEQ ID NO: 244	285 aa	MW at 31715.2kD
NOV38i, CG97012-04 Protein Sequence	CGGAVHNATIGRVLSPSPYENTNGSQFCIWTIEAPEGQKLHLHFERLLLHDKDRMTVHSG QTNKSALLYDSLQTESVPFEGLLSEGNTIRIEFTSDQARAASFTNIRFEAFKGHCYEPY IQNGNFTTSDPTYNIGTIVEFTCDPGHSLEQGPALIECINVRDPYWNDEPLCRAMCGGE LSAVAGVVLSPNWPEPYVEGEDCIWKIHVGEEKRIFLDIQFLNLSNSDILTIYDGDEVMP HILGQYLGNSGPQKLYSSTPDLTIQFHSDPAGLIFGKGQGFIMNY		
	SEQ ID NO: 245	2861 bp	
NOV38j, CG97012-05 DNA Sequence	AGCCACGATGCCCGCGGCCCGGCCCGCGCGGACTCCGCGGGATCTCGTGTTCCT CGCTCTGCTCCTGGGGAGCCCGCGCGCAGCGCTGGAGCGAGATGCTCTTCCCGAGGGAGA TGCTAGCCCTTTGGGTCTTACCTCCTGCCCTCAGGAGCCCCGAGAGAGGCAGTCTGG CAAAGAGCACCTGAAGAGAGAGTGGTAACAGCGCCCCCAGTTCCTCACAGTCGGCGGA AGTGCTGGGCGAGCTGGTGTGGATGGGACCGCACCTCTGCACATCAGCATACCCAGC CCTGTACCGCTGCTTCCAGAGGAGGCCCGCCCCAAGCACGCTTGGCCCCAAGAAGAA ACTGCCTTCGCTCAAGCAGGTGAACCTCTGCCAAGAGCAGCTGAGGCCCAAGGCCACCTC		

	CGCAGCCACTGTCCAAAGGGCAGGGTCCCAGCCAGCGTCCCAGGGCCTAGATCTCCTCTC CTCCTCCACGGAGAAGCCTGGCCCCACCGGGGACCCGGACCCCATCGTGGCCTCCGAGGA GGCATCAGAAAGTGGCCCTTTGGCTGGATCGAAAGGAGAGTGGGTCCCTACAACACCCGC ACCCCTGCAAATCTCCCCCTTCACTTCGCAGCCCTATGTGGCCACACACTCCCCAGAG GCCAGAACCCGGGAGCCTGGGCCTGACATGGCCAGGAGGCCCCAGGAGGACACCAG CCCCATGGCCCTGATGGACAAAGGTGAGAATGAGCTGACTGGGTGAGCCTCAGAGGAGAG CCAGGAGACCCTACCTCCACCATTATCACCACCAGGTGATCACCACCAGGACAGCACC AGCTCTCTGCAGTGTGAGCTTCTCCAATCCTGAGGGGTACATTGACTCCAGCGACTACCC ACTGCTGCCCCCTCAACAACCTTTCTGGAGTGCACATACAACGTGACAGTCTACACTGGCTA TGGGGTGGAGCTCCAGGTGAAGAGTGTGAACCTGTCCGATGGGGAACGTCTCCATCCG CGGGGTGGACGGCCCTACCTGACCGTCTGGCCAAACCAGACACTCCTGGTGGAGGGGCA GGTAATCCGAAGCCCCACCAACACCATCTCCGTCTACTTCCGGACCTTCCAGGACGACGG CCTTGGGACCTTCCAGCTTCACTACCAGGCCTTCACTGCTGAGCTGCACTTTCCCCCGCG GCCTGACTCTGGGGATGTACGGTGTGAGCCTGCACTCAGGTGGGGTGGCCCACTTTCA CTGCCACCTGGGCTATGAGCTCCAGGGCGCTAAGATGCTGACATGACATGCACTGCCAA GCCGCACTGGAGCAGCCAGGAGCCCATCTGCTCAGCTCCTTGTGGAGGGCAGTGCACAA TGCCACCATCGGCCGCTCTCTCCCAAGTTACCCTGAAAACACCAATGGGAGCCAATT CTGCATCTGGACGATTGAAGCTCCAGAGGGCCAGAAGCTGCACCTGCATTTGAGAGGCT GTTGCTGCATGACAAGGACAGGATGACGGTTACAGCGGGCAGACCAACAAGTCAGCTCT TCTCTACGACTCCCTTCAAACCGAGAGTGTCCCTTTGAGGGCCTGCTGAGCGAAGGCAA CACCATCCGCACTCGAGTTCAGTCCGACCGAGGCCGGCGCCTCCACCTTCAACATCCG ATTTGAAGCGTTTGAAGAGGCCACTGCTATGAGCCCTACATCCAGAATGGGAACCTTAC TACATCCGACCCGACCTATAACATTGGGACTATAGTGGAGTTCACCTGCCACCCCGGCA CTCCCTGGAGCAGGGCCCCGGCCATCATCGAATGCATCAATGTGCGGGACCCATACTGGAA TGACACAGAGCCCTGTGCAGAGCCATGTGTGGTGGGAGCTCTCTGCTGTGGCTGGGGT GGTATTGTCCCCAACTGGCCCCGAGCCCTACGTGGAAGGTGAAGATTGTATCTGGAAGAT CCACGTGGGAGAAGAGAAACGGATCTTCTTAGATATCCAGTTCCTGAATCTGAGCAACAG TGACATCTTGACCATCTACGATGGCGACGAGGTGATGCCCCACATCTTGGGGCAGTACCT TGGGAACAGTGGCCCCCAGAACTGTACTCCTCCACGCCAGACTTAACCATCCAGTTCCA TTCGGACCCTGCTGGCCTCATCTTTGGAAAGGGCCAGGGATTATCATGAACCTACATAGA GGTATCAAGGAATGACTCCTGCTCGGATTACCCGAGATCCAGAATGGCTGGAAAAACCAC TTCTCACACGGAGTTGGTGCGGGGAGCCAGAATCACCTACCAGTGTGACCCCGCTATGA CATCGTGGGGAGTGACACCCTCACCTGCCAGTGGGACCTCAGCTGGAGCAGCAGCCCCC ATTTTGTGAGAAAACGGAGGAGTCCCTGGCATGTGACAACCCAGGGCTGCTGAAAATGG ATACCAAATCCTGTACAAGCGACTCTACCTGCCAGGAGAGTCCCTCACCTTCATGTGCTA CGAAGGCTTTGAGCTCATGGGTGAAGTGACCATCCGCTGCATCTGGGACAGCCATCCA CTGGAACGGGCCCTGCCGTGTGTAAAGTAGCAGAAGCGGCAGCAGAGACGTGCTGGA AGGGGGGAACATGGCCCTGGCTATCTTCATCCCGTCTCATCATCTCCTTACTGTGGG AGGAGCCTACATTACATCACAAGATGTGCTACTATTCAAACCTCCGCTGCCTCTGAT GTACTCCCACCCCTACAGCCAGATCACCGTGGAAACCGAGTTTGACAACCCATTACGA GACAGGGGAAACCAGAGAGTATGAGGTTTCTATCTAAAGAG		
	ORF Start: ATG at 8		ORF Stop: TAA at 2855
	SEQ ID NO: 246	949 aa	MW at 103496.0kD
NOV38], CG97012-05 Protein Sequence	MPAARPPAAGLRGISLFLALLLGSAAAALERDALPEGDASPLGPYLLPSGAPERGSPGKE HPEERVVTAPPSSSQSAEVLGELVLDGTAPSAHHDIPALSPLLPEEARPKHALPPKKLP SLKQVNSARKQLRPKATSAAVQRAGSQPASQGLDLLSSSTEKPGPPGDPDIVASEEAS EVPLWLDRKESAVPTTPAPLQISPFSSQPYVAHTLPQRPEPEGEPGPDMAQEAPQEDTSPM ALMDKGENELTGSASEESQETTTSTIIITTTVITTEQAPALCSVSFSNPEGYIDSSDYPLL PLNNFLECTYNVTYVTGYGVELQVKS VNLSDGELLSIRGVDGPTLTVLANQTL LVEGQVI RSPTNTISVYFRTFQDDGLGTFQLHYQAFMLSCNFPRRPDSDGVTVMDLHSGGVAHFHCH LGYELQGAKMLTCLNASKPHWSSQEPICSAPCGAVHNATIGRVLSPSPYENTNGSQFCI WTIEAPEGQKLHLHFERLLHDKDRMTVHSGQTNKSALLYDSLQTESVPFEGLLSEGNTI RIEFTSDQARAASFNIRFEAFEGHCHYEPYIQNGNFTTSDPTYNIGTIVEFTCDPGHSL EQGPAILIECINVRDPYWNDEPLCRAMCGGELSAVAGVVLSPNWPEPYVEGEDCIWKIHV GEEKRIFLDIQFLNLSNSDILTIYDGDEVMPHILGQYLGNSGPQKLYSSTPDLTIQFHS PAGLIFGKGQGFIMNYIEVSRNDSCLDLPEIQNGWKTTSHTELVRGARITYQCDPGYDIV GSDTLTCQWDLWSDDPPFCEKTEESLACDNPLPENGQYILYKRLYLPGESLTFMCEYEG FELMGEVTIRCILGQPSHWNGPLPVCKVAEAAAETSLEGGNMALAFIPVLIISLLLGA		

YIYITRCRYYSNLRPLMYSHYPYSQITVETFDNPIYETGETREYEVS

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 38B.

Table 38B. Comparison of NOV38a against NOV38b through NOV38j.		
Protein Sequence	NOV38a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV38b	228..1013 1..797	751/797 (94%) 752/797 (94%)
NOV38c	393..865 1..474	427/477 (89%) 439/477 (91%)
NOV38d	30..1013 30..1013	944/984 (95%) 944/984 (95%)
NOV38e	452..738 3..289	285/287 (99%) 287/287 (99%)
NOV38f	452..738 3..289	285/287 (99%) 287/287 (99%)
NOV38g	452..738 3..289	284/287 (98%) 287/287 (99%)
NOV38h	392..866 2..477	429/479 (89%) 441/479 (91%)
NOV38i	452..736 1..285	285/285 (100%) 285/285 (100%)
NOV38j	30..872 30..873	752/847 (88%) 765/847 (89%)

Two polymorphic variants of NOV38a have been identified and are shown in Table 41P. Further analysis of the NOV38a protein yielded the following properties shown in Table 38C.

Table 38C. Protein Sequence Properties NOV38a	
PSort analysis:	0.6760 probability located in plasma membrane; 0.1800 probability located in nucleus; 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 29 and 30

A search of the NOV38a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 38D.

Table 38D. Geneseq Results for NOV38a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV38a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU12271	Human PRO6094 polypeptide sequence - <i>Homo sapiens</i> , 1023 aa. [WO200140466-A2, 07- JUN-2001]	1..1013 1..1023	1013/1023 (99%) 1013/1023 (99%)	0.0
ABG22405	Novel human diagnostic protein #22396 - <i>Homo sapiens</i> , 990 aa. [WO200175067-A2, 11- OCT-2001]	29..1013 6..990	983/985 (99%) 984/985 (99%)	0.0
ABG05922	Novel human diagnostic protein #5913 - <i>Homo sapiens</i> , 990 aa. [WO200175067-A2, 11- OCT-2001]	29..1013 6..990	983/985 (99%) 984/985 (99%)	0.0
ABG01221	Novel human diagnostic protein #1212 - <i>Homo sapiens</i> , 982 aa. [WO200175067-A2, 11- OCT-2001]	33..1013 2..982	981/981 (100%) 981/981 (100%)	0.0
ABG22407	Novel human diagnostic protein #22398 - <i>Homo sapiens</i> , 997 aa. [WO200175067-A2, 11- OCT-2001]	29..1008 6..996	967/991 (97%) 971/991 (97%)	0.0

In a BLAST search of public sequence databases, the NOV38a protein was found to have homology to the proteins shown in the BLASTP data in Table 38E.

Table 38E. Public BLASTP Results for NOV38a				
Protein Accession Number	Protein/Organism/Length	NOV38a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9BYH1	Seizure 6-like protein precursor - <i>Homo sapiens</i> (Human), 1024 aa.	1..1013 1..1024	1013/1024 (98%) 1013/1024 (98%)	0.0

Q9Y2E1	KIAA0927 protein - <i>Homo sapiens</i> (Human), 1001 aa (fragment).	1..872 53..925	821/876 (93%) 834/876 (94%)	0.0
Q9Y3J6	Hypothetical 87.6 kDa protein (DJ268D13.1.2) (seizure related gene 6 (mouse)-like (KIAA0927) (isoform 2)) - <i>Homo sapiens</i> (Human), 792 aa.	228..1008 1..791	778/791 (98%) 780/791 (98%)	0.0
Q9NUI3	DJ268D13.1.3 (Seizure related gene 6 (Mouse)-like (KIAA0927) (Isoform 3)) - <i>Homo sapiens</i> (Human), 777 aa (fragment).	228..1004 1..777	775/779 (99%) 775/779 (99%)	0.0
O95917	Hypothetical 79.0 kDa protein (DJ268D13.1.1) (seizure related gene 6 (mouse)-like (KIAA0927) (isoform 1)) - <i>Homo sapiens</i> (Human), 716 aa.	228..868 1..641	641/641 (100%) 641/641 (100%)	0.0

PFam analysis predicts that the NOV38a protein contains the domains shown in Table 38F.

Table 38F. Domain Analysis of NOV38a			
Pfam Domain	NOV38a Match Region	Identities/ Similarities for the Matched Region	Expect Value
sushi	393..448	16/65 (25%) 41/65 (63%)	6e-06
CUB	452..559	29/120 (24%) 72/120 (60%)	5.5e-09
sushi	567..624	19/67 (28%) 44/67 (66%)	4.5e-06
CUB	628..736	34/121 (28%) 69/121 (57%)	1.6e-15
sushi	745..800	22/64 (34%) 44/64 (69%)	1.3e-14
sushi	806..865	21/66 (32%) 47/66 (71%)	3.2e-11
sushi	873..930	20/65 (31%) 47/65 (72%)	4e-12

Example 39.

The NOV39 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 39A.

Table 39A. NOV39 Sequence Analysis			
	SEQ ID NO: 247	1957 bp	
NOV39a, CG99754-01 DNA Sequence	CAGGTGAGCAAGAGGATGCTGGCGGGGGCGTGAGGAGCATGCCAGCCCCCTCCTGGCC TGCTGGCAGCCCATCCTCCTGCTGGTGCTGGGCTCAGTGCTGTGAGGCTCGGCCACGGGC TGCCCGCCCCGCTGCGAGTGCTCCGCCAGGACCGCGCTGTGCTGTGCCACCGCAAGCGC TTTGTGGCAGTCCCGAGGGCATCCCCACCGAGACGCGCTGTGGACCTAGGCAAGAAC CGCATCAAAACGCTCAACCAGGACGAGTTCGCCAGCTTCCCGCACCTGGAGGAGCTGGAG CTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGCGCCTTCAACAACCTCTTCAACCTC CGGACGCTGGGTCTCCGACGCAACCGCCTGAAGCTCATCCCGCTAGGCGTCTTCACTGGC CTCAGCAACCTGACCAAGCTGGACATCAGCGAGAACAAGATCGTTATCCTACTGGACTAC ATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCGACAATGACCTCGTCTAC ATCTCTCACCGCGCTTCAAGCGCTTCAACAGCCTGGAGCAGCTGACGCTGGAGAAATGC AACCTGACCTCCATCCCCACCGAGGCGCTGTCCACCTGCACGGCCTCATCGTCTGAGG CTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGCTGTACCGACTC AAGGTCTTGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAACCTGCCTTAC GGCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCCCTACCTG GCCGTCCGCCACCTAGTCTATCTCCGCTTCTCAACCTCTCCTACAACCCATCAGCACC ATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGTGGTGGGCGGG CAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGCTCAACTACCTGCGCGTGCTCAAT GTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTG GAGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTCCGCTCCTGTGGGTGTT CGGCGCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCACGTGCGCCACGCCGAGTTT GTCCAGGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACTGC CGCCGCGCCCGCATCCGGGACCGCAAGGCCAGCAGGTGTTTGTGGACGAGGGCCACAG GTGCAGTTTGTGTGCGGGCGATGGCGACCCGCGCCCGCCATCCTCTGGCTCTCACCC CGAAAGCACCTGGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCAGC CTGGAGGTGCGCTACGCCAGGTACAGGACAACGGCAGTACCTGTGCTATCGCGGCCAAC GCGGGCGGCAACGACTCCATGCCCGCCACCTGCATGTGCGCAGTACTCGCCGACTGG CCCCATCAGCCCAACAAGACCTTCGCTTTCATCTCAACAGCCGGGCGAGGGAGAGGCC AACAGCACCCGCGCACTGTGCCTTTCCTTCGACATCAAGACCTCATCATCGCCACC ACCATGGGCTTCATCTCTTCTGGGCGTGTCTCTTCTGCCTGGTGTGCTGTTTCTC TGGAGCGGGGCAAGGGCAACAAGCACAAACATCGAGATCGAGTATGTGCCCGCAAG TCGGACGCAGGCATCAGCTCCGCCGACGCGCCCGCAAGTTCAACATGAAGATGATATGA GGCCGGGGCGGGGGCAGGACCCCCGGGCGGCGGGCAGGGGAAGGGGCTGGCCGCCA CCTGCTCACTCTCCAGTCTTCCACCTCCTCCCTAC		
	ORF Start: ATG at 16		ORF Stop: TGA at 1858
	SEQ ID NO: 248	614 aa	MW at 69145.1kD
NOV39a, CG99754-01 Protein Sequence	MLAGGVRSMPSPLLACWQPIILLVLGSLVLSGSATGCPPRCECSAQDRAVLCHRKRFAVP EGIPTETRLLDLGNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFNNLFLNRLTGL RSNRLKLIPLGVFTGLSNLTKLDISENKIVILLDMFQDLYNLKSLEVGDNDLVYISHRA FSGLNSLEQLTLEKCNLTSTPTEALSHLHGLIVLRHLNINAIKDYSPKRLYRLKVLKLEI SHWPYLDTMTNCLYGLNLTSLSTHNLTAVPYLAVRHLVLRFLNLSYNPISTIEGSM LHELLRLQEIQLVGGQLAVVEPYAFRLNLYLRLNVSGNQLTTLEESVFHVSNGNLETLLI DSNPLACDCRLLWVFRRLNFRNQPTCATPEFVQGEKDFPDVLLPNYFTCRRARI RDRKAQQVFVDEGHTVQFVCRADGPPAILWLSPRKHLVSAKSNRGLTVFPDGTLEVRY AQVQDNGTYLCAANAGGNDSPAHLHVRYSYSPDWPHQPNKTFAFISNQPGEGEANSTRA TVPFPFDIKTLIIATTMGFISFLGVLFCLVLLFLWSRGKNTKHNIEIYVPRKSDAGI SSADAPRKFNMKMI		
	SEQ ID NO: 249	2015 bp	
NOV39b,	GAGCTGAGGCTGGTGGGGGCGTGAGGAGCATGCCAGCCCCCTCCTGGCCTGCTGGCAG		

CG99754-02 DNA Sequence	CCCATCCTCCTGCTGGTGCTGGGCTCAGTGCTGTGTCAGGCTCGGGCCACGGGCTGCCCGCCC CGCTGCGAGTGCTCCGCCCAGGACCGCGCTGTGCTGTGCCACCGCAAGCGCTTTGTGGCA GTCCCCGAGGGCATCCCCACCGAGACGCGCTGTGACCTAGGCAAGAACCGCATCAAA ACGCTCAACCAGGACGAGTTCCGCCAGCTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAG AACATCGTGAGCGCCGTGGAGCCCGCGCCTTCAACAACCTCTTCAACCTCCGGACGCTG GGTCTCCGCAGCAACCGCCTGAAGCTCATCCCGCTAGGCGTCTTCACTGGCCTCAGAAC CTGACCAAGCTGGACATCAGCGAGAACAAGATCGTTATCCTACTGGACTACATGTTTCAG GACCTGTACAACCTCAAGTCACTGGAGGTTGGCGACAATGACCTCGTCTACATCTCTCAC CGCGCCTTACGCGGCCTCAACAGCCTGGAGCAGCTGACGCTGGAGAAATGCAACCTGACC TCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCCTCATCGTCTGAGGCTCCGGCAC CTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGCTGTACCGACTCAAGGTCTTG GAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAACCTGCCTCTACGGCCTCAAC CTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCCCTACCTGGCCGTCCGC CACCTAGTCTATCTCGCTTCTCTCAACCTCTCTACAACCCCATCAGCACCATTGAGGGC TCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGCGGGCAGCTGGCC GTGGTGGAGCCCTATGCCTTCCGCGGCCTCAACTACCTGCGCTGCTCAATGTCTCTGGC AACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGGAGACACTC ATCCTGGACTCCAACCGCTGGCCTGCGACTGTGCGTCTGTGGGTGTTCCGGCGCCGC TGGCGGCTCAACTTCAACCGGCAGCAGCCACGTGCGCCACGCCGAGTTTGTCCAGGGC AAGGAGTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACTGCCGCGCGGCC CGCATCCGGGACCGCAAGGCCACGAGGTGTTTGTGGACGAGGGCCACACGCTGCAGTTT GTGTGCCGGCCGATGGCGACCCGCGCCCGCCATCCTCTGGCTCTACCCCGAAAGCAC CTGGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCTCTGATGGCAGCTGGAGGTG CGCTACGCCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACCGGGCGGC AACGACTCCATGCCCGCCACCTGCATGTGCGCAGCTACTCGCCGACTGGCCCCATCAG CCCAACAGACCTTCGCTTTCATCTCAACCAGCCGGCGAGGAGAGGCCAACAGCACCC CGCGCCACTGTGCCTTTCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGC TTCATCTCTTCTGGGCGTCGCTCTTCTGCTGCTGCTGTTTCTCTGGAGCCGG GGCAAGGGCAACACAAAGCACACATCGAGATCGAGTATGTGCCCAAAAGTCGGACGCA GGCATCAGCTCCGCCGACGCGCCCGCAAGTTCAACATGAAGATGATAGAGGCCGGGGC GGGGGGCAGGGACCCCGGGCGCCGGGCAGGGGAAGGGGCTGGCCGCCACCTGCTCAC TCTCCAGTCTTCCACCTCTTCCCTACCTTCTACACAGTTCTCTTCTCCCTCCCGC CTCCGTCCCTGCTGCCCCCACCAGCCTCAGCTC		
	ORF Start: ATG at 31		ORF Stop: TGA at 1849
	SEQ ID NO: 250	606 aa	MW at 68345.1kD
NOV39b, CG99754-02 Protein Sequence	MPSPLLACWQPIILLVLGSLVSGSATGCPPRCECSAQDRAVLCHRRKFVAVPEGIPTETR LLDLGNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFFNNLFNRLTLGLRSNRLKLI PLGVFTGLSNLTKLDISENKIVILLDYMFDLYNLKSLEVGDNDLVYISHRAFSGLSLE QLTLEKCNLTISIPTALSHLHGLIVLRLRLNINAIIRDYSFKRLYLKVLKLEISHWPYLDT MTPNCLYGLNLTSLITHCNLTAVPYLAVRHLVYLRFLNLSYNPISTIEGSMHELLRLQ EIQLVGGQLAVVEPYAFRLNLYLRVLNVSGNQLTTLEESVFHVSNGLETILDSNPLACD CRLLWVFRRRWRLNFRNQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV FVDEGHTVQFVCRADGDPPIALWLSPRKHLVSAKSNGLTVFPDGTLEVRYAQVDNGT YLCIAANAGGNDSPAHLHVRSYSPDWPHPNKTFAFISNQPGEGEANSTRATVPFPFDI KTLIIATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPQKSDAGISSADAPRK FNMKMI		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 39B.

Table 39B. Comparison of NOV39a against NOV39b.		
Protein Sequence	NOV39a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV39b	9..614 1..606	563/606 (92%) 564/606 (92%)

Six polymorphic variants of NOV39a have been identified and are shown in Table 41Q. Further analysis of the NOV39a protein yielded the following properties shown in Table 39C.

Table 39C. Protein Sequence Properties NOV39a	
PSort analysis:	0.4600 probability located in plasma membrane; 0.1071 probability located in microbody (peroxisome); 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 36 and 37

- A search of the NOV39a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 39D.

Table 39D. Geneseq Results for NOV39a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV39a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB74705	Human membrane associated protein MEMAP-11 - <i>Homo sapiens</i> , 620 aa. [WO200112662-A2, 22-FEB-2001]	1..614 7..620	614/614 (100%) 614/614 (100%)	0.0
AAW84596	Amino acid sequence of the human Tango-79 protein - <i>Homo sapiens</i> , 614 aa. [WO9906427-A1, 11-FEB-1999]	1..614 1..614	612/614 (99%) 612/614 (99%)	0.0
AAB80225	Human PRO227 protein - <i>Homo sapiens</i> , 620 aa. [WO200104311-A1, 18-JAN-2001]	1..614 7..620	612/614 (99%) 612/614 (99%)	0.0
AAU12333	Human PRO227 polypeptide sequence - <i>Homo sapiens</i> , 620 aa. [WO200140466-A2, 07-JUN-2001]	1..614 7..620	612/614 (99%) 612/614 (99%)	0.0
AAI13357	Amino acid sequence of protein PRO227 - <i>Homo sapiens</i> , 620 aa. [WO9914328-A2, 25-MAR-1999]	1..614 7..620	612/614 (99%) 612/614 (99%)	0.0

In a BLAST search of public sequence databases, the NOV39a protein was found to have homology to the proteins shown in the BLASTP data in Table 39E.

Table 39E. Public BLASTP Results for NOV39a				
Protein Accession Number	Protein/Organism/Length	NOV39a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96FE5	Unknown (protein for MGC:17422) - <i>Homo sapiens</i> (Human), 614 aa.	1..614 1..614	614/614 (100%) 614/614 (100%)	0.0
Q9N008	Hypothetical 69.2 kDa protein - <i>Macaca fascicularis</i> (Crab eating macaque) (Cynomolgus monkey), 614 aa.	1..614 1..614	612/614 (99%) 613/614 (99%)	0.0
Q9D1T0	Adult male testis cDNA, RIKEN full-length enriched library, clone:4930471K13, full insert sequence - <i>Mus musculus</i> (Mouse), 614 aa.	1..614 1..614	610/614 (99%) 611/614 (99%)	0.0
CAD38935	Hypothetical protein - <i>Homo sapiens</i> (Human), 577 aa (fragment).	38..614 1..577	577/577 (100%) 577/577 (100%)	0.0
Q9BZ20	BA438B23.1 (Neuronal leucine-rich repeat protein) (CDNA FLJ31810 fis, clone NT2RI2009289, weakly similar to carboxypeptidase N 83 kDa chain) - <i>Homo sapiens</i> (Human), 606 aa.	14..614 6..606	365/603 (60%) 468/603 (77%)	0.0

PFam analysis predicts that the NOV39a protein contains the domains shown in

5 Table 39F.

Table 39F. Domain Analysis of NOV39a			
Pfam Domain	NOV39a Match Region	Identities/ Similarities for the Matched Region	Expect Value
LRRNT	35..64	10/31 (32%) 22/31 (71%)	0.00079

LRR	114..137	9/25 (36%) 20/25 (80%)	0.061
LRR	186..209	10/25 (40%) 19/25 (76%)	0.012
LRR	282..305	7/25 (28%) 17/25 (68%)	0.72
LRR	330..353	7/25 (28%) 20/25 (80%)	0.19
LRRCT	363..416	17/59 (29%) 39/59 (66%)	0.0021
ig	433..493	15/64 (23%) 44/64 (69%)	2.1e-09

Example 40.

The NOV40 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 40A.

Table 40A. NOV40 Sequence Analysis			
	SEQ ID NO: 251	889 bp	
NOV40a, CG99777-01 DNA Sequence	GGGAGAATCCTTCTTGAACAGAGATGGGCCAGAACTGAATCAGATGAAGAGAGATAAG GTGTGATGTGGGAAGACTATATAAAGAATGGACCCAGGGCTGCAGCAAGCACTCAACGG AATGGCCCTCCTGGAGACACAGCCATGCATGTGCCGGCGGGCTCCGTGGCCAGCCACCT GGGGACCACGAGCCGAGCTATTCTATTGACCACAGCCACTCTGGCTCTGTGCCTTGT CTTCACGGTGGCCACTATTATGGTGTGGTTCGTTTCAGAGGACGGACTCCATTCCCAACTC ACCTGACAACGTCCCCCTCAAAGGAGGAAATGCTCAGAAGACCTCTTATGTATCCTGAA AAGGGCTCCATTCAAGAAGTCATGGGCCTACCTCCAAGTGGCAAAGCATCTAAACAAAAC CAAGTTGTCTTGAACAAAGATGGCATTCTCCATGGAGTCAGATATCAGGATGGGAATCT GGTGATCCAATTCCTGGTTTGTACTTCATCATTTGCCAACTGCAGTTTCTGTACAATG CCCAAATAATTCTGTGATCTGAAGTTGGAGCTTCTCATCAACAGCATATCAAAAAACA GGCCCTGGTGACAGTGTGTGAGTCTGGAATGCAACGAAACACGTATACCAGAATCTCTC TCAATTCTGTGCTGGATTACCTGCAGGTCAACACCACCATATCAGTCAATGTGGATACATT CCAGTACATAGATACAAGCACCTTTCCTCTTGAGAATGTGTTGTCCATCTTCTTATACAG TAATTCAGACTGAACAGTTTCTCTTGGCCTTCAGGAAGAAAGCGCTCTCCACCATACAG TATTTTCATCCCTCCAAACACTTGGGCAAAAAGAAAACCTTTAGACCAAGA		
	ORF Start: ATG at 89		ORF Stop: TGA at 791
	SEQ ID NO: 252	234 aa	MW at 26016.9kD
NOV40a, CG99777-01 Protein Sequence	MDPGLQQALNGMAPPDGTAMHVPAGSVASHLGTTSRSYFYLTATLALCLVFTVATIMVL VVQRTDSIPNSPDNVPLKGGNCSDDLCLKRAPFKKSWAYLQVAKHLNKTLSWNKDIGI LHGVRYQDGNLVIQFPGLYFIICQLQLVQCPNNSVDLKLLELLINKHIKKQALVTVCESG MQTKHVVYQNLSQLLDYLVQNTTISVNVDTFQYIDTSTFPLENVLSIFLYSNSD		
	SEQ ID NO: 253	829 bp	
NOV40b, CG99777-02 DNA Sequence	GGGAGAATCCTTCTTGAACAGAGATGGGCCAGAACTGAATCAGATGAAGAGAGATAAG GTGTGATGTGGGAAGACTATATAAAGAATGGACCCAGGGCTGCAGCAAGCACTCAACGG AATGGCCCTCCTGGAGACACAGCCATGCATGTGCCGGCGGGCTCCGTGGCCAGCCACCT GGGGACCACGAGCCGAGCTATTCTATTGACCACAGCCACTCTGGCTCTGTGCCTTGT CTTCACGGTGGCCACTATTATGGTGTGGTTCGTTTCAGAGGACGGACTCCATTCCCAACTC ACCTGACAACGTCCCCCTCAAAGGAGTGGCAAAGCATCTAAACAAAACCAAGTTGTCTTG GAACAAAGATGGCATTCTCCATGGAGTCAGATATCAGGATGGGAATCTGGTGATCCAATT CCCTGGTTTGTACTTCATCATTTGCCAACTGCAGTTTCTTGTAATGCCAAATAATTCT		

	TGTCGATCTGAAGTTGGAGCTTCTCATCAACAAGCATATCAAAAAACAGGCCCTGGTGAC AGTGTGTGAGTCTGGAATGCAAACGAAACACGTATACCAGAATCTCTCTCAATTCTTGCT GGATTACCTGCAGGTCAACACCACCATATCAGTCAATGTGGATACATTCCAGTACATAGA TACAAGCACCTTTCTCTTGAGAATGTGTTGTCCATCTTCTATACAGTAATTTCAGACTG AACAGTTTCTCTTGGCCTTCAGGAAGAAAGCGCCTCTCTACCATACAGTATTTTCATCCCT CCAAACACTTGGGCAAAAAGAAACTTTAGACCAAGAAGGATTCTCTC		
	ORF Start: ATG at 89		ORF Stop: TGA at 719
	SEQ ID NO: 254	210 aa	MW at 23250.6kD
NOV40b, CG99777-02 Protein Sequence	MDPGLQQALNGMAPPGDTAMHVPAGSVASHLGTTSRSYFYLTATLALCLVFTVATIMVL VVQRTDSIPNSPDNVPLKGVAKHLNKTLSWNKDGI LHGVRYQDGNLVIQFPGLYFIICQ LQFLVQCPNNSVDLKLELLINKHIKKQALVTVCESGMQTKHVYQNL SQFLLDYLQVNTTI SVNVDTFQYIDTSTFPLENVLSIFLYSNSD		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 40B.

Table 40B. Comparison of NOV40a against NOV40b.		
Protein Sequence	NOV40a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV40b	1..234 1..210	210/234 (89%) 210/234 (89%)

Three polymorphic variants of NOV40b have been identified and are shown in Table 41R.

- 5 Further analysis of the NOV40a protein yielded the following properties shown in Table 40C.

Table 40C. Protein Sequence Properties NOV40a	
PSort analysis:	0.7900 probability located in plasma membrane; 0.3000 probability located in microbody (peroxisome); 0.3000 probability located in Golgi body; 0.2000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	Cleavage site between residues 68 and 69

A search of the NOV40a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 40D.

10

Table 40D. Geneseq Results for NOV40a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV40a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value

AAU78086	Human CD30-ligand (CD30L) protein sequence - <i>Homo sapiens</i> , 234 aa. [WO200211767-A2, 14-FEB-2002]	1..234 1..234	234/234 (100%) 234/234 (100%)	e-135
AAR45009	Sequence encoded by a human CD30-L cDNA clone encoding additional N-terminal amino acids - <i>Homo sapiens</i> , 234 aa. [WO9324135-A, 09-DEC-1993]	1..234 1..234	234/234 (100%) 234/234 (100%)	e-135
AAR45007	Sequence encoded by a human CD30-L cDNA clone - <i>Homo sapiens</i> , 215 aa. [WO9324135-A, 09-DEC-1993]	20..234 1..215	215/215 (100%) 215/215 (100%)	e-123
AAU78087	Mouse CD30-ligand (CD30L) protein sequence - <i>Mus</i> sp, 239 aa. [WO200211767-A2, 14-FEB-2002]	1..234 1..239	167/240 (69%) 195/240 (80%)	4e-92
AAR45008	Sequence encoded by a murine CD30-L cDNA clone encoding additional N-terminal amino acids - <i>Acomys cahirinus</i> , 239 aa. [WO9324135-A, 09-DEC-1993]	1..234 1..239	167/240 (69%) 195/240 (80%)	4e-92

In a BLAST search of public sequence databases, the NOV40a protein was found to have homology to the proteins shown in the BLASTP data in Table 40E.

Table 40E. Public BLASTP Results for NOV40a				
Protein Accession Number	Protein/Organism/Length	NOV40a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P32971	Tumor necrosis factor ligand superfamily member 8 (CD30 ligand) (CD30- L) (CD153 antigen) - <i>Homo sapiens</i> (Human), 234 aa.	1..234 1..234	234/234 (100%) 234/234 (100%)	e-134

P32972	Tumor necrosis factor ligand superfamily member 8 (CD30 ligand) (CD30- L) - <i>Mus musculus</i> (Mouse), 239 aa.	1..234 1..239	167/240 (69%) 195/240 (80%)	1e-91
AAD46392	CD30 LIGAND-EXOTOXIN A FUSION PROTEIN - synthetic construct, 220 aa (fragment).	86..234 48..196	149/149 (100%) 149/149 (100%)	9e-83
P41047	Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand) - <i>Mus musculus</i> (Mouse), 279 aa.	97..195 142..264	31/123 (25%) 53/123 (42%)	0.056
Q9WV90	Fas ligand - <i>Marmota monax</i> (Woodchuck), 169 aa (fragment).	100..154 44..101	20/58 (34%) 29/58 (49%)	0.49

PFam analysis predicts that the NOV40a protein contains the domains shown in Table 40F.

Table 40F. Domain Analysis of NOV40a			
Pfam Domain	NOV40a Match Region	Identities/ Similarities for the Matched Region	Expect Value
TNF	93..230	55/159 (35%) 136/159 (86%)	1.6e-53

Example B: Sequencing Methodology and Identification of NOVX Clones

- 5 1. **GeneCalling™ Technology:** This is a proprietary method of performing differential gene expression profiling between two or more samples developed at CuraGen and described by Shimkets, *et al.*, "Gene expression analysis by transcript profiling coupled to a gene database query" Nature Biotechnology 17:198-803 (1999). cDNA was derived from various human samples representing multiple tissue types, normal and diseased states,
- 10 physiological states, and developmental states from different donors. Samples were obtained as whole tissue, primary cells or tissue cultured primary cells or cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression, for example, growth factors, chemokines or steroids. The cDNA thus derived was then digested with up to as many as 120 pairs of restriction enzymes and pairs of
- 15 linker-adaptors specific for each pair of restriction enzymes were ligated to the appropriate

end. The restriction digestion generates a mixture of unique cDNA gene fragments. Limited PCR amplification is performed with primers homologous to the linker adapter sequence where one primer is biotinylated and the other is fluorescently labeled. The doubly labeled material is isolated and the fluorescently labeled single strand is resolved by capillary gel electrophoresis. A computer algorithm compares the electropherograms from an experimental and control group for each of the restriction digestions. This and additional sequence-derived information is used to predict the identity of each differentially expressed gene fragment using a variety of genetic databases. The identity of the gene fragment is confirmed by additional, gene-specific competitive PCR or by isolation and sequencing of the gene fragment.

2. **SeqCalling™ Technology:** cDNA was derived from various human samples representing multiple tissue types, normal and diseased states, physiological states, and developmental states from different donors. Samples were obtained as whole tissue, primary cells or tissue cultured primary cells or cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression, for example, growth factors, chemokines or steroids. The cDNA thus derived was then sequenced using CuraGen's proprietary SeqCalling technology. Sequence traces were evaluated manually and edited for corrections if appropriate. cDNA sequences from all samples were assembled together, sometimes including public human sequences, using bioinformatic programs to produce a consensus sequence for each assembly. Each assembly is included in CuraGen Corporation's database. Sequences were included as components for assembly when the extent of identity with another component was at least 95% over 50 bp. Each assembly represents a gene or portion thereof and includes information on variants, such as splice forms single nucleotide polymorphisms (SNPs), insertions, deletions and other sequence variations.

3. **PathCalling™ Technology:** The NOVX nucleic acid sequences are derived by laboratory screening of cDNA library by the two-hybrid approach. cDNA fragments covering either the full length of the DNA sequence, or part of the sequence, or both, are sequenced. In silico prediction was based on sequences available in CuraGen Corporation's proprietary sequence databases or in the public human sequence databases, and provided either the full length DNA sequence, or some portion thereof.

The laboratory screening was performed using the methods summarized below:

cDNA libraries were derived from various human samples representing multiple tissue types, normal and diseased states, physiological states, and developmental states from different donors. Samples were obtained as whole tissue, primary cells or tissue cultured primary cells or cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression, for example, growth factors, chemokines or steroids. The cDNA thus derived was then directionally cloned into the appropriate two-hybrid vector (Gal4-activation domain (Gal4-AD) fusion). Such cDNA libraries as well as commercially available cDNA libraries from Clontech (Palo Alto, CA) were then transferred from E.coli into a CuraGen Corporation proprietary yeast strain (disclosed in U. S. Patents 6,057,101 and 6,083,693, incorporated herein by reference in their entireties).

Gal4-binding domain (Gal4-BD) fusions of a CuraGen Corporation proprietary library of human sequences was used to screen multiple Gal4-AD fusion cDNA libraries resulting in the selection of yeast hybrid diploids in each of which the Gal4-AD fusion contains an individual cDNA. Each sample was amplified using the polymerase chain reaction (PCR) using non-specific primers at the cDNA insert boundaries. Such PCR product was sequenced; sequence traces were evaluated manually and edited for corrections if appropriate. cDNA sequences from all samples were assembled together, sometimes including public human sequences, using bioinformatic programs to produce a consensus sequence for each assembly. Each assembly is included in CuraGen Corporation's database. Sequences were included as components for assembly when the extent of identity with another component was at least 95% over 50 bp. Each assembly represents a gene or portion thereof and includes information on variants, such as splice forms single nucleotide polymorphisms (SNPs), insertions, deletions and other sequence variations.

Physical clone: the cDNA fragment derived by the screening procedure, covering the entire open reading frame is, as a recombinant DNA, cloned into pACT2 plasmid (Clontech) used to make the cDNA library. The recombinant plasmid is inserted into the host and selected by the yeast hybrid diploid generated during the screening procedure by the mating of both CuraGen Corporation proprietary yeast strains N106' and YULH (U. S. Patents 6,057,101 and 6,083,693).

4. **RACE:** Techniques based on the polymerase chain reaction such as rapid amplification of cDNA ends (RACE), were used to isolate or complete the predicted sequence of the cDNA of the invention. Usually multiple clones were sequenced from one

or more human samples to derive the sequences for fragments. Various human tissue samples from different donors were used for the RACE reaction. The sequences derived from these procedures were included in the SeqCalling Assembly process described in preceding paragraphs.

- 5 **5. Exon Linking:** The NOVX target sequences identified in the present invention were subjected to the exon linking process to confirm the sequence. PCR primers were designed by starting at the most upstream sequence available, for the forward primer, and at the most downstream sequence available for the reverse primer. In each case, the sequence was examined, walking inward from the respective termini toward the coding
- 10 sequence, until a suitable sequence that is either unique or highly selective was encountered, or, in the case of the reverse primer, until the stop codon was reached. Such primers were designed based on in silico predictions for the full length cDNA, part (one or more exons) of the DNA or protein sequence of the target sequence, or by translated homology of the predicted exons to closely related human sequences from other species. These primers were
- 15 then employed in PCR amplification based on the following pool of human cDNAs: adrenal gland, bone marrow, brain - amygdala, brain - cerebellum, brain - hippocampus, brain - substantia nigra, brain - thalamus, brain -whole, fetal brain, fetal kidney, fetal liver, fetal lung, heart, kidney, lymphoma - Raji, mammary gland, pancreas, pituitary gland, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord, spleen,
- 20 stomach, testis, thyroid, trachea, uterus. Usually the resulting amplicons were gel purified, cloned and sequenced to high redundancy. The PCR product derived from exon linking was cloned into the pCR2.1 vector from Invitrogen. The resulting bacterial clone has an insert covering the entire open reading frame cloned into the pCR2.1 vector. The resulting sequences from all clones were assembled with themselves, with other fragments in
- 25 CuraGen Corporation's database and with public ESTs. Fragments and ESTs were included as components for an assembly when the extent of their identity with another component of the assembly was at least 95% over 50 bp. In addition, sequence traces were evaluated manually and edited for corrections if appropriate. These procedures provide the sequence reported herein.
- 30 **6. Physical Clone:** Exons were predicted by homology and the intron/exon boundaries were determined using standard genetic rules. Exons were further selected and refined by means of similarity determination using multiple BLAST (for example, tBlastN,

BlastX, and BlastN) searches, and, in some instances, GeneScan and Grail. Expressed sequences from both public and proprietary databases were also added when available to further define and complete the gene sequence. The DNA sequence was then manually corrected for apparent inconsistencies thereby obtaining the sequences encoding the full-length protein.

The PCR product derived by exon linking, covering the entire open reading frame, was cloned into the pCR2.1 vector from Invitrogen to provide clones used for expression and screening purposes.

Example C: Quantitative expression analysis of clones in various cells and tissues

The quantitative expression of various clones was assessed using microtiter plates containing RNA samples from a variety of normal and pathology-derived cells, cell lines and tissues using real time quantitative PCR (RTQ PCR). RTQ PCR was performed on an Applied Biosystems ABI PRISM® 7700 or an ABI PRISM® 7900 HT Sequence Detection System. Various collections of samples are assembled on the plates, and referred to as Panel 1 (containing normal tissues and cancer cell lines), Panel 2 (containing samples derived from tissues from normal and cancer sources), Panel 3 (containing cancer cell lines), Panel 4 (containing cells and cell lines from normal tissues and cells related to inflammatory conditions), Panel 5D/5I (containing human tissues and cell lines with an emphasis on metabolic diseases), AI_comprehensive_panel (containing normal tissue and samples from autoimmune/autoinflammatory diseases), Panel CNSD.01 (containing samples from normal and diseased brains) and CNS_neurodegeneration_panel (containing samples from normal and Alzheimer's diseased brains).

RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase using probe and primer sets designed to amplify across the span of a single exon.

First, the RNA samples were normalized to reference nucleic acids such as constitutively expressed genes (for example, β -actin and GAPDH). Normalized RNA (5 ul)

was converted to cDNA and analyzed by RTQ-PCR using One Step RT-PCR Master Mix Reagents (Applied Biosystems; Catalog No. 4309169) and gene-specific primers according to the manufacturer's instructions.

In other cases, non-normalized RNA samples were converted to single strand cDNA (sscDNA) using Superscript II (Invitrogen Corporation; Catalog No. 18064-147) and random hexamers according to the manufacturer's instructions. Reactions containing up to 10 µg of total RNA were performed in a volume of 20 µl and incubated for 60 minutes at 42 °C. This reaction can be scaled up to 50 µg of total RNA in a final volume of 100 µl. sscDNA samples are then normalized to reference nucleic acids as described previously, using 1X TaqMan® Universal Master mix (Applied Biosystems; catalog No. 4324020), following the manufacturer's instructions.

Probes and primers were designed for each assay according to Applied Biosystems Primer Express Software package (version I for Apple Computer's Macintosh Power PC) or a similar algorithm using the target sequence as input. Default settings were used for reaction conditions and the following parameters were set before selecting primers: primer concentration = 250 nM, primer melting temperature (T_m) range = 58 °-60 °C, primer optimal T_m = 59 °C, maximum primer difference = 2 °C, probe does not have 5'G, probe T_m must be 10 °C greater than primer T_m, amplicon size 75bp to 100bp. The probes and primers selected (see below) were synthesized by SyntheGen (Houston, TX, USA). Probes were double purified by HPLC to remove uncoupled dye and evaluated by mass spectroscopy to verify coupling of reporter and quencher dyes to the 5' and 3' ends of the probe, respectively. Their final concentrations were: forward and reverse primers, 900 nM each, and probe, 200 nM.

PCR conditions: When working with RNA samples, normalized RNA from each tissue and each cell line was spotted in each well of either a 96 well or a 384-well PCR plate (Applied Biosystems). PCR cocktails included either a single gene specific probe and primers set, or two multiplexed probe and primers sets (a set specific for the target clone and another gene-specific set multiplexed with the target probe). PCR reactions were set up using TaqMan® One-Step RT-PCR Master Mix (Applied Biosystems, Catalog No. 4313803) following manufacturer's instructions. Reverse transcription was performed at 48°C for 30 minutes followed by amplification/PCR cycles as follows: 95°C 10 min, then 40 cycles of 95 °C for 15 seconds, 60 °C for 1 minute. Results were recorded as CT values

(cycle at which a given sample crosses a threshold level of fluorescence) using a log scale, with the difference in RNA concentration between a given sample and the sample with the lowest CT value being represented as 2 to the power of delta CT. The percent relative expression is then obtained by taking the reciprocal of this RNA difference and multiplying
5 by 100.

When working with sscDNA samples, normalized sscDNA was used as described previously for RNA samples. PCR reactions containing one or two sets of probe and primers were set up as described previously, using 1X TaqMan® Universal Master mix (Applied Biosystems; catalog No. 4324020), following the manufacturer's instructions. PCR
10 amplification was performed as follows: 95 °C 10 min, then 40 cycles of 95 °C for 15 seconds, 60 °C for 1 minute. Results were analyzed and processed as described previously.

Panels 1, 1.1, 1.2, and 1.3D

The plates for Panels 1, 1.1, 1.2 and 1.3D include 2 control wells (genomic DNA control and chemistry control) and 94 wells containing cDNA from various samples. The
15 samples in these panels are broken into 2 classes: samples derived from cultured cell lines and samples derived from primary normal tissues. The cell lines are derived from cancers of the following types: lung cancer, breast cancer, melanoma, colon cancer, prostate cancer, CNS cancer, squamous cell carcinoma, ovarian cancer, liver cancer, renal cancer, gastric cancer and pancreatic cancer. Cell lines used in these panels are widely available through
20 the American Type Culture Collection (ATCC), a repository for cultured cell lines, and were cultured using the conditions recommended by the ATCC. The normal tissues found on these panels are comprised of samples derived from all major organ systems from single adult individuals or fetuses. These samples are derived from the following organs: adult skeletal muscle, fetal skeletal muscle, adult heart, fetal heart, adult kidney, fetal kidney,
25 adult liver, fetal liver, adult lung, fetal lung, various regions of the brain, the spleen, bone marrow, lymph node, pancreas, salivary gland, pituitary gland, adrenal gland, spinal cord, thymus, stomach, small intestine, colon, bladder, trachea, breast, ovary, uterus, placenta, prostate, testis and adipose.

In the results for Panels 1, 1.1, 1.2 and 1.3D, the following abbreviations are used:

30 ca. = carcinoma,

* = established from metastasis,

met = metastasis,

s cell var = small cell variant,

non-s = non-sm = non-small,

squam = squamous,

5 pl. eff = pl effusion = pleural effusion,

glio = glioma,

astro = astrocytoma, and

neuro = neuroblastoma.

General_screening_panel_v1.4, v1.5 and v1.6

10 The plates for Panels 1.4, v1.5 and v1.6 include two control wells (genomic DNA control and chemistry control) and 94 wells containing cDNA from various samples. The samples in Panels 1.4, v1.5 and v1.6 are broken into 2 classes: samples derived from cultured cell lines and samples derived from primary normal tissues. The cell lines are derived from cancers of the following types: lung cancer, breast cancer, melanoma, colon
15 cancer, prostate cancer, CNS cancer, squamous cell carcinoma, ovarian cancer, liver cancer, renal cancer, gastric cancer and pancreatic cancer. Cell lines used in Panels 1.4, v1.5 and v1.6 are widely available through the American Type Culture Collection (ATCC), a repository for cultured cell lines, and were cultured using the conditions recommended by the ATCC. The normal tissues found on Panels 1.4, v1.5 and v1.6 are comprised of pools of
20 samples derived from all major organ systems from 2 to 5 different adult individuals or fetuses. These samples are derived from the following organs: adult skeletal muscle, fetal skeletal muscle, adult heart, fetal heart, adult kidney, fetal kidney, adult liver, fetal liver, adult lung, fetal lung, various regions of the brain, the spleen, bone marrow, lymph node, pancreas, salivary gland, pituitary gland, adrenal gland, spinal cord, thymus, stomach, small
25 intestine, colon, bladder, trachea, breast, ovary, uterus, placenta, prostate, testis and adipose. Abbreviations are as described for Panels 1, 1.1, 1.2, and 1.3D.

Panels 2D, 2.2, 2.3 and 2.4

The plates for Panels 2D, 2.2, 2.3 and 2.4 generally include two control wells and 94 test samples composed of RNA or cDNA isolated from human tissue procured by surgeons
30 working in close cooperation with the National Cancer Institute's Cooperative Human

Tissue Network (CHTN) or the National Disease Research Initiative (NDRI) or from Ardaïs or Clinomics. The tissues are derived from human malignancies and in cases where indicated many malignant tissues have "matched margins" obtained from noncancerous tissue just adjacent to the tumor. These are termed normal adjacent tissues and are denoted "NAT" in the results below. The tumor tissue and the "matched margins" are evaluated by two independent pathologists (the surgical pathologists and again by a pathologist at NDRI/CHTN/Ardaïs/Clinomics). Unmatched RNA samples from tissues without malignancy (normal tissues) were also obtained from Ardaïs or Clinomics. This analysis provides a gross histopathological assessment of tumor differentiation grade. Moreover, most samples include the original surgical pathology report that provides information regarding the clinical stage of the patient. These matched margins are taken from the tissue surrounding (*i.e.* immediately proximal) to the zone of surgery (designated "NAT", for normal adjacent tissue, in Table RR). In addition, RNA and cDNA samples were obtained from various human tissues derived from autopsies performed on elderly people or sudden death victims (accidents, *etc.*). These tissues were ascertained to be free of disease and were purchased from various commercial sources such as Clontech (Palo Alto, CA), Research Genetics, and Invitrogen. General oncology screening panel_v_2.4 is an updated version of Panel 2D.

HASS Panel v 1.0

The HASS panel v 1.0 plates are comprised of 93 cDNA samples and two controls. Specifically, 81 of these samples are derived from cultured human cancer cell lines that had been subjected to serum starvation, acidosis and anoxia for different time periods as well as controls for these treatments, 3 samples of human primary cells, 9 samples of malignant brain cancer (4 medulloblastomas and 5 glioblastomas) and 2 controls. The human cancer cell lines are obtained from ATCC (American Type Culture Collection) and fall into the following tissue groups: breast cancer, prostate cancer, bladder carcinomas, pancreatic cancers and CNS cancer cell lines. These cancer cells are all cultured under standard recommended conditions. The treatments used (serum starvation, acidosis and anoxia) have been previously published in the scientific literature. The primary human cells were obtained from Clonetics (Walkersville, MD) and were grown in the media and conditions recommended by Clonetics. The malignant brain cancer samples are obtained as part of a collaboration (Henry Ford Cancer Center) and are evaluated by a pathologist prior to

CuraGen receiving the samples. RNA was prepared from these samples using the standard procedures. The genomic and chemistry control wells have been described previously.

ARDAIS Panel v 1.0

The plates for ARDAIS panel v 1.0 generally include 2 control wells and 22 test
5 samples composed of RNA isolated from human tissue procured by surgeons working in
close cooperation with Ardais Corporation. The tissues are derived from human lung
malignancies (lung adenocarcinoma or lung squamous cell carcinoma) and in cases where
indicated many malignant samples have “matched margins” obtained from noncancerous
lung tissue just adjacent to the tumor. These matched margins are taken from the tissue
10 surrounding (*i.e.* immediately proximal) to the zone of surgery (designated “NAT”, for
normal adjacent tissue) in the results below. The tumor tissue and the “matched margins”
are evaluated by independent pathologists (the surgical pathologists and again by a
pathologist at Ardais). Unmatched malignant and non-malignant RNA samples from lungs
were also obtained from Ardais. Additional information from Ardais provides a gross
15 histopathological assessment of tumor differentiation grade and stage. Moreover, most
samples include the original surgical pathology report that provides information regarding
the clinical state of the patient.

Panels 3D and 3.1

The plates of Panels 3D and 3.1 are comprised of 94 cDNA samples and two control
20 samples. Specifically, 92 of these samples are derived from cultured human cancer cell
lines, 2 samples of human primary cerebellar tissue and 2 controls. The human cell lines are
generally obtained from ATCC (American Type Culture Collection), NCI or the German
tumor cell bank and fall into the following tissue groups: Squamous cell carcinoma of the
tongue, breast cancer, prostate cancer, melanoma, epidermoid carcinoma, sarcomas, bladder
25 carcinomas, pancreatic cancers, kidney cancers, leukemias/lymphomas,
ovarian/uterine/cervical, gastric, colon, lung and CNS cancer cell lines. In addition, there
are two independent samples of cerebellum. These cells are all cultured under standard
recommended conditions and RNA extracted using the standard procedures. The cell lines
in panel 3D and 1.3D are of the most common cell lines used in the scientific literature.
30 Oncology_cell_line_screening_panel_v3.2 is an updated version of Panel 3. The cell lines
in panel 3D, 3.1, 1.3D and oncology_cell_line_screening_panel_v3.2 are of the most
common cell lines used in the scientific literature.

Panels 4D, 4R, and 4.1D

Panel 4 includes samples on a 96 well plate (2 control wells, 94 test samples) composed of RNA (Panel 4R) or cDNA (Panels 4D/4.1D) isolated from various human cell lines or tissues related to inflammatory conditions. Total RNA from control normal tissues such as colon and lung (Stratagene, La Jolla, CA) and thymus and kidney (Clontech) was employed. Total RNA from liver tissue from cirrhosis patients and kidney from lupus patients was obtained from BioChain (Biochain Institute, Inc., Hayward, CA). Intestinal tissue for RNA preparation from patients diagnosed as having Crohn's disease and ulcerative colitis was obtained from the National Disease Research Interchange (NDRI) (Philadelphia, PA).

Astrocytes, lung fibroblasts, dermal fibroblasts, coronary artery smooth muscle cells, small airway epithelium, bronchial epithelium, microvascular dermal endothelial cells, microvascular lung endothelial cells, human pulmonary aortic endothelial cells, human umbilical vein endothelial cells were all purchased from Clonetics (Walkersville, MD) and grown in the media supplied for these cell types by Clonetics. These primary cell types were activated with various cytokines or combinations of cytokines for 6 and/or 12-14 hours, as indicated. The following cytokines were used; IL-1 beta at approximately 1-5ng/ml, TNF alpha at approximately 5-10ng/ml, IFN gamma at approximately 20-50ng/ml, IL-4 at approximately 5-10ng/ml, IL-9 at approximately 5-10ng/ml, IL-13 at approximately 5-10ng/ml. Endothelial cells were sometimes starved for various times by culture in the basal media from Clonetics with 0.1% serum.

Mononuclear cells were prepared from blood of employees at CuraGen Corporation, using Ficoll. LAK cells were prepared from these cells by culture in DMEM 5% FCS (Hyclone), 100µM non essential amino acids (Gibco/Life Technologies, Rockville, MD), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10 mM Hepes (Gibco) and Interleukin 2 for 4-6 days. Cells were then either activated with 10-20ng/ml PMA and 1-2µg/ml ionomycin, IL-12 at 5-10ng/ml, IFN gamma at 20-50ng/ml and IL-18 at 5-10ng/ml for 6 hours. In some cases, mononuclear cells were cultured for 4-5 days in DMEM 5% FCS (Hyclone), 100 µM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10 mM Hepes (Gibco) with PHA (phytohemagglutinin) or PWM (pokeweed mitogen) at approximately 5 µg/ml. Samples were taken at 24, 48 and 72 hours for RNA preparation. MLR (mixed lymphocyte

reaction) samples were obtained by taking blood from two donors, isolating the mononuclear cells using Ficoll and mixing the isolated mononuclear cells 1:1 at a final concentration of approximately 2×10^6 cells/ml in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol
5 (5.5 $\times 10^{-5}$ M) (Gibco), and 10 mM Hepes (Gibco). The MLR was cultured and samples taken at various time points ranging from 1- 7 days for RNA preparation.

Monocytes were isolated from mononuclear cells using CD14 Miltenyi Beads, +ve VS selection columns and a Vario Magnet according to the manufacturer's instructions. Monocytes were differentiated into dendritic cells by culture in DMEM 5% fetal calf serum
10 (FCS) (Hyclone, Logan, UT), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 $\times 10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco), 50ng/ml GMCSF and 5ng/ml IL-4 for 5-7 days. Macrophages were prepared by culture of monocytes for 5-7 days in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 $\times 10^{-5}$ M (Gibco), 10 mM
15 Hepes (Gibco) and 10% AB Human Serum or MCSF at approximately 50ng/ml. Monocytes, macrophages and dendritic cells were stimulated for 6 and 12-14 hours with lipopolysaccharide (LPS) at 100ng/ml. Dendritic cells were also stimulated with anti-CD40 monoclonal antibody (Pharmingen) at 10 μ g/ml for 6 and 12-14 hours.

CD4 lymphocytes, CD8 lymphocytes and NK cells were also isolated from
20 mononuclear cells using CD4, CD8 and CD56 Miltenyi beads, positive VS selection columns and a Vario Magnet according to the manufacturer's instructions. CD45RA and CD45RO CD4 lymphocytes were isolated by depleting mononuclear cells of CD8, CD56, CD14 and CD19 cells using CD8, CD56, CD14 and CD19 Miltenyi beads and positive selection. CD45RO beads were then used to isolate the CD45RO CD4 lymphocytes with the
25 remaining cells being CD45RA CD4 lymphocytes. CD45RA CD4, CD45RO CD4 and CD8 lymphocytes were placed in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 $\times 10^{-5}$ M (Gibco), and 10 mM Hepes (Gibco) and plated at 10^6 cells/ml onto Falcon 6 well tissue culture plates that had been coated overnight with 0.5 μ g/ml anti-CD28 (Pharmingen) and 3 μ g/ml anti-CD3
30 (OKT3, ATCC) in PBS. After 6 and 24 hours, the cells were harvested for RNA preparation. To prepare chronically activated CD8 lymphocytes, we activated the isolated CD8 lymphocytes for 4 days on anti-CD28 and anti-CD3 coated plates and then harvested

the cells and expanded them in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10 mM Hepes (Gibco) and IL-2. The expanded CD8 cells were then activated again with plate bound anti-CD3 and anti-CD28 for 4 days and expanded as before. RNA was isolated 6 and 24 hours after the second activation and after 4 days of the second expansion culture. The isolated NK cells were cultured in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10 mM Hepes (Gibco) and IL-2 for 4-6 days before RNA was prepared.

To obtain B cells, tonsils were procured from NDRI. The tonsil was cut up with sterile dissecting scissors and then passed through a sieve. Tonsil cells were then spun down and resuspended at 10^6 cells/ml in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10 mM Hepes (Gibco). To activate the cells, we used PWM at 5 μ g/ml or anti-CD40 (Pharmingen) at approximately 10 μ g/ml and IL-4 at 5-10ng/ml. Cells were harvested for RNA preparation at 24, 48 and 72 hours.

To prepare the primary and secondary Th1/Th2 and Tr1 cells, six-well Falcon plates were coated overnight with 10 μ g/ml anti-CD28 (Pharmingen) and 2 μ g/ml OKT3 (ATCC), and then washed twice with PBS. Umbilical cord blood CD4 lymphocytes (Poietic Systems, German Town, MD) were cultured at 10^5 - 10^6 cells/ml in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), 10 mM Hepes (Gibco) and IL-2 (4ng/ml). IL-12 (5ng/ml) and anti-IL4 (1 μ g/ml) were used to direct to Th1, while IL-4 (5ng/ml) and anti-IFN gamma (1 μ g/ml) were used to direct to Th2 and IL-10 at 5ng/ml was used to direct to Tr1. After 4-5 days, the activated Th1, Th2 and Tr1 lymphocytes were washed once in DMEM and expanded for 4-7 days in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), 10 mM Hepes (Gibco) and IL-2 (1ng/ml). Following this, the activated Th1, Th2 and Tr1 lymphocytes were re-stimulated for 5 days with anti-CD28/OKT3 and cytokines as described above, but with the addition of anti-CD95L (1 μ g/ml) to prevent apoptosis. After 4-5 days, the Th1, Th2 and Tr1 lymphocytes were washed and then expanded again with IL-2 for 4-7 days. Activated Th1 and Th2 lymphocytes were maintained in this way for a maximum of three cycles. RNA was prepared from primary and secondary Th1, Th2 and Tr1 after 6 and 24 hours

following the second and third activations with plate bound anti-CD3 and anti-CD28 mAbs and 4 days into the second and third expansion cultures in Interleukin 2.

The following leukocyte cells lines were obtained from the ATCC: Ramos, EOL-1, KU-812. EOL cells were further differentiated by culture in 0.1 mM dbcAMP at
 5 5×10^5 cells/ml for 8 days, changing the media every 3 days and adjusting the cell concentration to 5×10^5 cells/ml. For the culture of these cells, we used DMEM or RPMI (as recommended by the ATCC), with the addition of 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), 10 mM Hepes (Gibco). RNA was either prepared from resting cells or cells
 10 activated with PMA at 10 ng/ml and ionomycin at 1 μ g/ml for 6 and 14 hours. Keratinocyte line CCD106 and an airway epithelial tumor line NCI-H292 were also obtained from the ATCC. Both were cultured in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10 mM Hepes (Gibco). CCD106 cells were activated for 6 and 14 hours with approximately 5
 15 ng/ml TNF alpha and 1 ng/ml IL-1 beta, while NCI-H292 cells were activated for 6 and 14 hours with the following cytokines: 5 ng/ml IL-4, 5 ng/ml IL-9, 5 ng/ml IL-13 and 25 ng/ml IFN gamma.

For these cell lines and blood cells, RNA was prepared by lysing approximately 10^7 cells/ml using Trizol (Gibco BRL). Briefly, 1/10 volume of bromochloropropane
 20 (Molecular Research Corporation) was added to the RNA sample, vortexed and after 10 minutes at room temperature, the tubes were spun at 14,000 rpm in a Sorvall SS34 rotor. The aqueous phase was removed and placed in a 15 ml Falcon Tube. An equal volume of isopropanol was added and left at -20°C overnight. The precipitated RNA was spun down at 9,000 rpm for 15 min in a Sorvall SS34 rotor and washed in 70% ethanol. The pellet was
 25 redissolved in 300 μ l of RNase-free water and 35 μ l buffer (Promega) 5 μ l DTT, 7 μ l RNasin and 8 μ l DNase were added. The tube was incubated at 37°C for 30 minutes to remove contaminating genomic DNA, extracted once with phenol chloroform and re-precipitated with 1/10 volume of 3M sodium acetate and 2 volumes of 100% ethanol. The RNA was spun down and placed in RNase free water. RNA was stored at -80°C .

30 AI_comprehensive panel_v1.0

The plates for AI_comprehensive panel_v1.0 include two control wells and 89 test samples comprised of cDNA isolated from surgical and postmortem human tissues obtained

from the Backus Hospital and Clinomics (Frederick, MD). Total RNA was extracted from tissue samples from the Backus Hospital in the Facility at CuraGen. Total RNA from other tissues was obtained from Clinomics.

5 Joint tissues including synovial fluid, synovium, bone and cartilage were obtained from patients undergoing total knee or hip replacement surgery at the Backus Hospital. Tissue samples were immediately snap frozen in liquid nitrogen to ensure that isolated RNA was of optimal quality and not degraded. Additional samples of osteoarthritis and rheumatoid arthritis joint tissues were obtained from Clinomics. Normal control tissues were supplied by Clinomics and were obtained during autopsy of trauma victims.

10 Surgical specimens of psoriatic tissues and adjacent matched tissues were provided as total RNA by Clinomics. Two male and two female patients were selected between the ages of 25 and 47. None of the patients were taking prescription drugs at the time samples were isolated.

15 Surgical specimens of diseased colon from patients with ulcerative colitis and Crohn's disease and adjacent matched tissues were obtained from Clinomics. Bowel tissue from three female and three male Crohn's patients between the ages of 41-69 were used. Two patients were not on prescription medication while the others were taking dexamethasone, phenobarbital, or tylenol. Ulcerative colitis tissue was from three male and four female patients. Four of the patients were taking lebid and two were on phenobarbital.

20 Total RNA from post mortem lung tissue from trauma victims with no disease or with emphysema, asthma or COPD was purchased from Clinomics. Emphysema patients ranged in age from 40-70 and all were smokers, this age range was chosen to focus on patients with cigarette-linked emphysema and to avoid those patients with alpha-1-anti-trypsin deficiencies. Asthma patients ranged in age from 36-75, and excluded
25 smokers to prevent those patients that could also have COPD. COPD patients ranged in age from 35-80 and included both smokers and non-smokers. Most patients were taking corticosteroids, and bronchodilators.

In the labels employed to identify tissues in the AI_comprehensive panel_v1.0 panel, the following abbreviations are used:

30 AI = Autoimmunity

Syn = Synovial

Normal = No apparent disease

Rep22 /Rep20 = individual patients

RA = Rheumatoid arthritis

Backus = From Backus Hospital

5 OA = Osteoarthritis

(SS) (BA) (MF) = Individual patients

Adj = Adjacent tissue

Match control = adjacent tissues

-M = Male

10 -F = Female

COPD = Chronic obstructive pulmonary disease

Panels 5D and 5I

The plates for Panel 5D and 5I include two control wells and a variety of cDNAs isolated from human tissues and cell lines with an emphasis on metabolic diseases.

15 Metabolic tissues were obtained from patients enrolled in the Gestational Diabetes study. Cells were obtained during different stages in the differentiation of adipocytes from human mesenchymal stem cells. Human pancreatic islets were also obtained.

In the Gestational Diabetes study subjects are young (18 - 40 years), otherwise healthy women with and without gestational diabetes undergoing routine (elective)
20 Caesarean section. After delivery of the infant, when the surgical incisions were being repaired/closed, the obstetrician removed a small sample (<1 cc) of the exposed metabolic tissues during the closure of each surgical level. The biopsy material was rinsed in sterile saline, blotted and fast frozen within 5 minutes from the time of removal. The tissue was then flash frozen in liquid nitrogen and stored, individually, in sterile screw-top tubes and
25 kept on dry ice for shipment to or to be picked up by CuraGen. The metabolic tissues of interest include uterine wall (smooth muscle), visceral adipose, skeletal muscle (rectus) and subcutaneous adipose. Patient descriptions are as follows:

Patient 2 Diabetic Hispanic, overweight, not on insulin

Patient 7-9 Nondiabetic Caucasian and obese (BMI>30)

Patient 10 Diabetic Hispanic, overweight, on insulin

Patient 11 Nondiabetic African American and overweight

Patient 12 Diabetic Hispanic on insulin

Adipocyte differentiation was induced in donor progenitor cells obtained from
 5 Osirus (a division of Clonetics/BioWhittaker) in triplicate, except for Donor 3U which had
 only two replicates. Scientists at Clonetics isolated, grew and differentiated human
 mesenchymal stem cells (HuMSCs) for CuraGen based on the published protocol found in
 Mark F. Pittenger, *et al.*, Multilineage Potential of Adult Human Mesenchymal Stem Cells
 Science Apr 2 1999: 143-147. Clonetics provided Trizol lysates or frozen pellets suitable
 10 for mRNA isolation and ds cDNA production. A general description of each donor is as
 follows:

Donor 2 and 3 U: Mesenchymal Stem cells, Undifferentiated Adipose

Donor 2 and 3 AM: Adipose, AdiposeMidway Differentiated

Donor 2 and 3 AD: Adipose, Adipose Differentiated

15 Human cell lines were generally obtained from ATCC (American Type Culture
 Collection), NCI or the German tumor cell bank and fall into the following tissue groups:
 kidney proximal convoluted tubule, uterine smooth muscle cells, small intestine, liver
 HepG2 cancer cells, heart primary stromal cells, and adrenal cortical adenoma cells. These
 cells are all cultured under standard recommended conditions and RNA extracted using the
 20 standard procedures. All samples were processed at CuraGen to produce single stranded
 cDNA.

Panel 5I contains all samples previously described with the addition of pancreatic
 islets from a 58 year old female patient obtained from the Diabetes Research Institute at the
 University of Miami School of Medicine. Islet tissue was processed to total RNA at an
 25 outside source and delivered to CuraGen for addition to panel 5I.

In the labels employed to identify tissues in the 5D and 5I panels, the following
 abbreviations are used:

GO Adipose = Greater Omentum Adipose

SK = Skeletal Muscle

UT = Uterus

PL = Placenta

AD = Adipose Differentiated

AM = Adipose Midway Differentiated

5 U = Undifferentiated Stem Cells

Panel CNSD.01

The plates for Panel CNSD.01 include two control wells and 94 test samples comprised of cDNA isolated from postmortem human brain tissue obtained from the Harvard Brain Tissue Resource Center. Brains are removed from calvaria of donors
10 between 4 and 24 hours after death, sectioned by neuroanatomists, and frozen at -80°C in liquid nitrogen vapor. All brains are sectioned and examined by neuropathologists to confirm diagnoses with clear associated neuropathology.

Disease diagnoses are taken from patient records. The panel contains two brains from each of the following diagnoses: Alzheimer's disease, Parkinson's disease,
15 Huntington's disease, Progressive Supranuclear Palsy, Depression, and "Normal controls". Within each of these brains, the following regions are represented: cingulate gyrus, temporal pole, globus pallidus, substantia nigra, Brodman Area 4 (primary motor strip), Brodman Area 7 (parietal cortex), Brodman Area 9 (prefrontal cortex), and Brodman area 17 (occipital cortex). Not all brain regions are represented in all cases; *e.g.*, Huntington's
20 disease is characterized in part by neurodegeneration in the globus pallidus, thus this region is impossible to obtain from confirmed Huntington's cases. Likewise Parkinson's disease is characterized by degeneration of the substantia nigra making this region more difficult to obtain. Normal control brains were examined for neuropathology and found to be free of any pathology consistent with neurodegeneration.

25 In the labels employed to identify tissues in the CNS panel, the following abbreviations are used:

PSP = Progressive supranuclear palsy

Sub Nigra = Substantia nigra

Glob Palladus= Globus pallidus

Temp Pole = Temporal pole

Cing Gyr = Cingulate gyrus

BA 4 = Brodman Area 4

Panel CNS_Neurodegeneration_V1.0

5 The plates for Panel CNS_Neurodegeneration_V1.0 include two control wells and
47 test samples comprised of cDNA isolated from postmortem human brain tissue obtained
from the Harvard Brain Tissue Resource Center (McLean Hospital) and the Human Brain
and Spinal Fluid Resource Center (VA Greater Los Angeles Healthcare System). Brains are
removed from calvaria of donors between 4 and 24 hours after death, sectioned by
10 neuroanatomists, and frozen at -80°C in liquid nitrogen vapor. All brains are sectioned and
examined by neuropathologists to confirm diagnoses with clear associated neuropathology.

Disease diagnoses are taken from patient records. The panel contains six brains from
Alzheimer's disease (AD) patients, and eight brains from "Normal controls" who showed no
evidence of dementia prior to death. The eight normal control brains are divided into two
15 categories: Controls with no dementia and no Alzheimer's like pathology (Controls) and
controls with no dementia but evidence of severe Alzheimer's like pathology, (specifically
senile plaque load rated as level 3 on a scale of 0-3; 0 = no evidence of plaques, 3 = severe
AD senile plaque load). Within each of these brains, the following regions are represented:
hippocampus, temporal cortex (Brodman Area 21), parietal cortex (Brodman area 7), and
20 occipital cortex (Brodman area 17). These regions were chosen to encompass all levels of
neurodegeneration in AD. The hippocampus is a region of early and severe neuronal loss in
AD; the temporal cortex is known to show neurodegeneration in AD after the hippocampus;
the parietal cortex shows moderate neuronal death in the late stages of the disease; the
occipital cortex is spared in AD and therefore acts as a "control" region within AD patients.
25 Not all brain regions are represented in all cases.

In the labels employed to identify tissues in the CNS_Neurodegeneration_V1.0
panel, the following abbreviations are used:

AD = Alzheimer's disease brain; patient was demented and showed AD-like
pathology upon autopsy

30 Control = Control brains; patient not demented, showing no neuropathology

Control (Path) = Control brains; pateint not demented but showing sever
AD-like pathology

SupTemporal Ctx = Superior Temporal Cortex

Inf Temporal Ctx = Inferior Temporal Cortex

5

A. CG133274-02: Induced Myeloid Leukemia Cell

Differentiation Protein MCL-1-like Protein.

Expression of gene CG133274-02 was assessed using the primer-probe set Ag7050,
described in Table AA. Results of the RTQ-PCR runs are shown in Table AB.

10 Table AA. Probe Name Ag7050

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gtctcgtggttgcgctg-3'	17	450	255
Probe	TET-5'- tcgtaagggtctccagcgccttcctg- 3'-TAMRA	25	485	256
Reverse	5'-gattggcgccaaggaca-3'	17	541	257

Table AB. General_screening_panel_v1.6

Tissue Name	Rel. Exp.(%) Ag7050, Run 282273858	Tissue Name	Rel. Exp.(%) Ag7050, Run 282273858
Adipose	100.0	Renal ca. TK-10	59.5
Melanoma* Hs688(A).T	33.7	Bladder	60.7
Melanoma* Hs688(B).T	34.9	Gastric ca. (liver met.) NCI-N87	87.1
Melanoma* M14	33.0	Gastric ca. KATO III	59.0
Melanoma* LOXIMV1	49.0	Colon ca. SW-948	19.8
Melanoma* SK-MEL-5	22.8	Colon ca. SW480	36.1
Squamous cell carcinoma SCC-4	19.2	Colon ca.* (SW480 met) SW620	25.0
Testis Pool	12.7	Colon ca. HT29	28.7
Prostate ca.* (bone met) PC-3	44.4	Colon ca. HCT-116	56.6
Prostate Pool	18.3	Colon ca. CaCo-2	24.0
Placenta	27.2	Colon cancer tissue	69.3
Uterus Pool	14.0	Colon ca. SW1116	12.1

Ovarian ca. OVCAR-3	46.3	Colon ca. Colo-205	10.2
Ovarian ca. SK-OV-3	53.6	Colon ca. SW-48	11.0
Ovarian ca. OVCAR-4	32.1	Colon Pool	14.4
Ovarian ca. OVCAR-5	55.5	Small Intestine Pool	24.5
Ovarian ca. IGROV-1	31.4	Stomach Pool	19.3
Ovarian ca. OVCAR-8	34.4	Bone Marrow Pool	11.5
Ovary	21.5	Fetal Heart	13.5
Breast ca. MCF-7	72.2	Heart Pool	13.9
Breast ca. MDA-MB-231	60.3	Lymph Node Pool	16.0
Breast ca. BT 549	81.2	Fetal Skeletal Muscle	8.5
Breast ca. T47D	18.2	Skeletal Muscle Pool	17.0
Breast ca. MDA-N	12.3	Spleen Pool	59.9
Breast Pool	14.6	Thymus Pool	24.0
Trachea	44.1	CNS cancer (glio/astro) U87-MG	82.4
Lung	11.5	CNS cancer (glio/astro) U-118-MG	42.3
Fetal Lung	81.2	CNS cancer (neuro;met) SK-N-AS	46.7
Lung ca. NCI-N417	15.3	CNS cancer (astro) SF- 539	22.1
Lung ca. LX-1	61.6	CNS cancer (astro) SNB- 75	45.4
Lung ca. NCI-H146	17.8	CNS cancer (glio) SNB- 19	31.9
Lung ca. SHP-77	55.9	CNS cancer (glio) SF-295	60.7
Lung ca. A549	28.1	Brain (Amygdala) Pool	6.9
Lung ca. NCI-H526	25.2	Brain (cerebellum)	12.7
Lung ca. NCI-H23	90.1	Brain (fetal)	10.0
Lung ca. NCI-H460	37.1	Brain (Hippocampus) Pool	10.2
Lung ca. HOP-62	27.2	Cerebral Cortex Pool	8.4
Lung ca. NCI-H522	33.9	Brain (Substantia nigra) Pool	7.0
Liver	2.7	Brain (Thalamus) Pool	7.6
Fetal Liver	14.1	Brain (whole)	4.5
Liver ca. HepG2	20.4	Spinal Cord Pool	22.8
Kidney Pool	40.9	Adrenal Gland	19.9
Fetal Kidney	11.3	Pituitary gland Pool	5.9
Renal ca. 786-0	31.6	Salivary Gland	7.3
Renal ca. A498	12.3	Thyroid (female)	30.6
Renal ca. ACHN	31.9	Pancreatic ca. CAPAN2	25.7
Renal ca. UO-31	33.9	Pancreas Pool	29.7

General_screening_panel_v1.6 Summary: Ag7050 Highest expression of this gene is seen in adipose (CT=25). This gene is ubiquitously expressed in this panel, with high to moderate expression seen in brain, colon, gastric, lung, breast, ovarian, and melanoma cancer cell lines. This expression profile suggests a role for this gene product in cell survival and proliferation. Modulation of this gene product may be useful in the treatment of cancer.

Among tissues with metabolic function, this gene is expressed at high to moderate levels in pituitary, adipose, adrenal gland, pancreas, thyroid, and adult and fetal skeletal muscle, heart, and liver. This widespread expression among these tissues suggests that this gene product may play a role in normal neuroendocrine and metabolic function and that disregulated expression of this gene may contribute to neuroendocrine disorders or metabolic diseases, such as obesity and diabetes.

This gene is also expressed at moderate levels in the CNS, including the hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurologic disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

B. CG134430-01: RIKEN cDNA 2310034L04 Like Gene.

Expression of gene CG134430-01 was assessed using the primer-probe set Ag7372, described in Table BA. Results of the RTQ-PCR runs are shown in Table BB.

Table BA. Probe Name Ag7372

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - catttgaagtgggtgtctacacttataaa -3'	28	789	258
Probe	TET-5' - agtctgtccctctgggtgttctcac- 3' -TAMRA'	26	818	259
Reverse	5' - ggcatagatatttctctgattacttcata t-3'	29	860	260

Table BB. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag7372, Run 305065597	Tissue Name	Rel. Exp.(%) Ag7372, Run 305065597
Secondary Th1 act	46.7	HUVEC IL-1beta	19.1
Secondary Th2 act	51.1	HUVEC IFN gamma	17.9
Secondary Tr1 act	27.0	HUVEC TNF alpha + IFN gamma	3.2
Secondary Th1 rest	6.1	HUVEC TNF alpha + IL4	3.9
Secondary Th2 rest	16.8	HUVEC IL-11	8.5
Secondary Tr1 rest	13.9	Lung Microvascular EC none	36.9
Primary Th1 act	25.2	Lung Microvascular EC TNFalpha + IL-1beta	8.1
Primary Th2 act	80.7	Microvascular Dermal EC none	3.6
Primary Tr1 act	58.2	Microvascular Dermal EC TNFalpha + IL-1beta	4.0
Primary Th1 rest	4.4	Bronchial epithelium TNFalpha + IL1beta	1.4
Primary Th2 rest	5.4	Small airway epithelium none	4.6
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	31.6
CD45RA CD4 lymphocyte act	100.0	Coronary artery SMC rest	6.0
CD45RO CD4 lymphocyte act	37.9	Coronary artery SMC TNFalpha + IL-1beta	10.4
CD8 lymphocyte act	18.2	Astrocytes rest	2.2
Secondary CD8 lymphocyte rest	3.5	Astrocytes TNFalpha + IL- 1beta	1.7
Secondary CD8 lymphocyte act	6.9	KU-812 (Basophil) rest	8.1
CD4 lymphocyte none	13.1	KU-812 (Basophil) PMA/ionomycin	11.3
2ry Th1/Th2/Tr1_anti- CD95 CH11	21.2	CCD1106 (Keratinocytes) none	18.3
LAK cells rest	5.1	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	4.6
LAK cells IL-2	41.8	Liver cirrhosis	3.9
LAK cells IL-2+IL-12	1.3	NCI-H292 none	8.4
LAK cells IL-2+IFN gamma	4.0	NCI-H292 IL-4	6.4
LAK cells IL-2+ IL-18	3.4	NCI-H292 IL-9	9.0
LAK cells PMA/ionomycin	36.3	NCI-H292 IL-13	7.5
NK Cells IL-2 rest	82.4	NCI-H292 IFN gamma	2.1
Two Way MLR 3 day	14.0	HPAEC none	7.4
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	32.1

Two Way MLR 7 day	9.1	Lung fibroblast none	11.9
PBMC rest	8.3	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	8.2	Lung fibroblast IL-4	9.5
PBMC PHA-L	6.6	Lung fibroblast IL-9	6.4
Ramos (B cell) none	9.1	Lung fibroblast IL-13	10.6
Ramos (B cell) ionomycin	30.1	Lung fibroblast IFN gamma	9.1
B lymphocytes PWM	9.5	Dermal fibroblast CCD1070 rest	25.7
B lymphocytes CD40L and IL-4	41.2	Dermal fibroblast CCD1070 TNF alpha	47.3
EOL-1 dbcAMP	29.7	Dermal fibroblast CCD1070 IL-1 beta	22.2
EOL-1 dbcAMP PMA/ionomycin	22.8	Dermal fibroblast IFN gamma	8.9
Dendritic cells none	13.8	Dermal fibroblast IL-4	11.8
Dendritic cells LPS	9.2	Dermal Fibroblasts rest	15.0
Dendritic cells anti-CD40	9.8	Neutrophils TNFa+LPS	31.2
Monocytes rest	13.1	Neutrophils rest	98.6
Monocytes LPS	16.2	Colon	5.4
Macrophages rest	4.9	Lung	4.5
Macrophages LPS	9.1	Thymus	16.2
HUVEC none	37.1	Kidney	30.8
HUVEC starved	10.8		

Panel 4.1D Summary: Ag7372 This gene is widely expressed at low levels in many samples on this panel. Highest expression of this gene is seen in CD45RA CD4 cells, naive T cells that have been activated with CD3 and CD28 (CT=32.6). Significant expression is also seen in both acutely and chronically activated T cells, resting neutrophils and NK cells. Based on the widespread expression of this gene in cells of significance to the autoimmune response, modulation of the expression or function of this gene may be useful in the treatment of autoimmune disease, including T cell mediated diseases such as asthma, arthritis, psoriasis, inflammatory bowel disease, and lupus.

C. CG137677-01 and CG137697-01: RIKEN 5730409G15-like protein.

Expression of gene CG137677-01 and CG137697-01 was assessed using the primer-probe sets Ag4928 and Ag4927, described in Tables CA and CB. Results of the RTQ-PCR runs are shown in Tables CC, CD and CE.

Table CA. Probe Name Ag4928

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tcagatgggaagtggaagct-3'	20	935	261
Probe	TET-5'- ccagaaactgtttccctacagagagca -3'-TAMRA	27	963	262
Reverse	5'-aggttcagcattgccatct-3'	19	995	263

Table CB. Probe Name Ag4927

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ccccaggcatacatcttca-3'	19	571	264
Probe	TET-5'- actgtcacagccgggtcctcgag- 3'-TAMRA	23	593	265
Reverse	5'-gaggccattgagaaggacat- 3'	20	629	266

Table CC. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4927, Run 224735008	Rel. Exp.(%) Ag4928, Run 224735009	Tissue Name	Rel. Exp.(%) Ag4927, Run 224735008	Rel. Exp.(%) Ag4928, Run 224735009
AD 1 Hippo	4.7	14.2	Control (Path) 3 Temporal Ctx	8.7	11.7
AD 2 Hippo	42.3	66.9	Control (Path) 4 Temporal Ctx	32.8	51.8
AD 3 Hippo	4.9	7.9	AD 1 Occipital Ctx	14.4	9.1
AD 4 Hippo	9.8	12.6	AD 2 Occipital Ctx (Missing)	0.0	0.0
AD 5 hippo	65.5	83.5	AD 3 Occipital Ctx	3.9	4.9
AD 6 Hippo	40.6	82.9	AD 4 Occipital Ctx	23.7	18.2
Control 2 Hippo	23.5	25.7	AD 5 Occipital Ctx	0.0	66.4
Control 4 Hippo	16.0	17.7	AD 6 Occipital Ctx	33.0	13.4
Control (Path) 3 Hippo	0.0	15.7	Control 1 Occipital Ctx	6.1	10.2
AD 1 Temporal Ctx	12.7	22.5	Control 2 Occipital Ctx	61.6	47.0
AD 2 Temporal Ctx	44.4	70.2	Control 3 Occipital Ctx	25.3	54.3
AD 3 Temporal Ctx	5.3	2.9	Control 4 Occipital Ctx	8.7	8.4

AD 4 Temporal Ctx	4.1	27.0	Control (Path) 1 Occipital Ctx	100.0	100.0
AD 5 Inf Temporal Ctx	55.9	86.5	Control (Path) 2 Occipital Ctx	4.2	18.3
AD 5 Sup Temporal Ctx	33.9	70.2	Control (Path) 3 Occipital Ctx	3.8	4.6
AD 6 Inf Temporal Ctx	38.4	40.3	Control (Path) 4 Occipital Ctx	27.5	31.9
AD 6 Sup Temporal Ctx	54.3	49.3	Control 1 Parietal Ctx	8.9	19.5
Control 1 Temporal Ctx	8.7	12.0	Control 2 Parietal Ctx	33.2	46.7
Control 2 Temporal Ctx	37.4	51.8	Control 3 Parietal Ctx	4.1	30.6
Control 3 Temporal Ctx	18.7	23.0	Control (Path) 1 Parietal Ctx	76.3	73.2
Control 4 Temporal Ctx	17.1	20.7	Control (Path) 2 Parietal Ctx	31.6	29.3
Control (Path) 1 Temporal Ctx	62.4	77.9	Control (Path) 3 Parietal Ctx	5.0	15.9
Control (Path) 2 Temporal Ctx	59.5	48.6	Control (Path) 4 Parietal Ctx	57.0	52.9

Table CD. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag4927, Run 228839257	Rel. Exp.(%) Ag4928, Run 228839262	Tissue Name	Rel. Exp.(%) Ag4927, Run 228839257	Rel. Exp.(%) Ag4928, Run 228839262
Adipose	1.4	3.1	Renal ca. TK-10	35.1	41.8
Melanoma* Hs688(A).T	12.9	18.4	Bladder	13.2	10.4
Melanoma* Hs688(B).T	12.7	18.4	Gastric ca. (liver met.) NCI-N87	82.4	98.6
Melanoma* M14	33.4	34.4	Gastric ca. KATO III	100.0	68.3
Melanoma* LOXIMVI	34.9	25.2	Colon ca. SW-948	15.8	10.2
Melanoma* SK-MEL-5	37.1	68.8	Colon ca. SW480	43.8	55.5
Squamous cell carcinoma SCC-4	21.8	21.5	Colon ca. * (SW480 met) SW620	41.8	44.4
Testis Pool	9.5	6.5	Colon ca. HT29	22.8	15.8
Prostate ca. * (bone met) PC-3	11.7	26.8	Colon ca. HCT-116	54.0	47.0

Prostate Pool	4.4	4.8	Colon ca. CaCo-2	31.0	47.3
Placenta	3.2	3.8	Colon cancer tissue	9.4	7.0
Uterus Pool	2.0	4.2	Colon ca. SW1116	8.4	6.1
Ovarian ca. OVCAR-3	25.7	34.9	Colon ca. Colo-205	5.9	7.3
Ovarian ca. SK-OV-3	20.6	19.5	Colon ca. SW-48	9.2	7.5
Ovarian ca. OVCAR-4	8.1	9.9	Colon Pool	17.3	8.5
Ovarian ca. OVCAR-5	44.8	44.8	Small Intestine Pool	17.7	10.8
Ovarian ca. IGROV-1	8.8	27.0	Stomach Pool	11.8	3.7
Ovarian ca. OVCAR-8	16.0	10.2	Bone Marrow Pool	5.2	3.2
Ovary	7.5	8.0	Fetal Heart	3.1	4.1
Breast ca. MCF-7	24.7	28.9	Heart Pool	4.6	3.7
Breast ca. MDA-MB-231	14.7	20.9	Lymph Node Pool	21.2	15.5
Breast ca. BT 549	24.3	12.8	Fetal Skeletal Muscle	5.8	5.0
Breast ca. T47D	8.7	11.1	Skeletal Muscle Pool	5.3	4.4
Breast ca. MDA-N	18.0	18.6	Spleen Pool	3.8	2.7
Breast Pool	18.6	8.7	Thymus Pool	15.9	9.5
Trachea	9.4	8.0	CNS cancer (glio/astro) U87-MG	35.1	49.3
Lung	5.3	4.5	CNS cancer (glio/astro) U-118-MG	40.9	40.6
Fetal Lung	17.3	13.7	CNS cancer (neuro;met) SK-N-AS	16.7	22.4
Lung ca. NCI-N417	8.2	5.7	CNS cancer (astro) SF-539	10.7	7.1
Lung ca. LX-1	70.2	77.4	CNS cancer (astro) SNB-75	27.4	17.9
Lung ca. NCI-H146	9.7	4.3	CNS cancer (glio) SNB-19	12.0	15.7
Lung ca. SHP-77	34.2	26.6	CNS cancer (glio) SF-295	38.2	43.2
Lung ca. A549	25.5	29.7	Brain (Amygdala) Pool	8.2	3.4
Lung ca. NCI-H526	3.1	5.4	Brain (cerebellum)	22.5	16.2
Lung ca. NCI-H23	43.5	100.0	Brain (fetal)	25.5	7.9

Lung ca. NCI-H460	25.2	24.1	Brain (Hippocampus) Pool	7.7	4.4
Lung ca. HOP-62	12.7	11.8	Cerebral Cortex Pool	10.5	5.0
Lung ca. NCI-H522	51.1	48.6	Brain (Substantia nigra) Pool	9.3	4.4
Liver	1.2	1.6	Brain (Thalamus) Pool	15.0	6.7
Fetal Liver	9.5	9.9	Brain (whole)	11.2	5.1
Liver ca. HepG2	22.1	23.2	Spinal Cord Pool	11.0	5.1
Kidney Pool	25.9	13.4	Adrenal Gland	10.7	8.9
Fetal Kidney	15.7	17.6	Pituitary gland Pool	6.7	9.6
Renal ca. 786-0	12.6	14.9	Salivary Gland	3.7	4.7
Renal ca. A498	7.9	7.0	Thyroid (female)	3.4	4.8
Renal ca. ACHN	27.2	23.8	Pancreatic ca. CAPAN2	37.4	31.9
Renal ca. UO-31	21.5	20.2	Pancreas Pool	23.0	13.3

Table CE. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4927, Run 223598856	Rel. Exp.(%) Ag4928, Run 223597247	Tissue Name	Rel. Exp.(%) Ag4927, Run 223598856	Rel. Exp.(%) Ag4928, Run 223597247
Secondary Th1 act	25.2	13.8	HUVEC IL-1beta	15.8	9.7
Secondary Th2 act	23.5	8.7	HUVEC IFN gamma	11.2	6.5
Secondary Tr1 act	12.5	8.9	HUVEC TNF alpha + IFN gamma	16.4	6.7
Secondary Th1 rest	6.6	4.2	HUVEC TNF alpha + IL4	20.9	10.8
Secondary Th2 rest	6.4	4.2	HUVEC IL-11	10.4	4.3
Secondary Tr1 rest	6.3	2.1	Lung Microvascular EC none	29.1	12.9
Primary Th1 act	31.0	18.2	Lung Microvascular EC TNFalpha + IL-1beta	27.0	13.0
Primary Th2 act	22.8	13.0	Microvascular Dermal EC none	11.2	2.5
Primary Tr1 act	29.9	15.9	Microvascular Dermal EC TNFalpha + IL-1beta	13.8	6.9
Primary Th1 rest	5.7	1.6	Bronchial epithelium TNFalpha + IL1beta	10.1	6.7

Primary Th2 rest	3.8	2.5	Small airway epithelium none	7.3	2.5
Primary Tr1 rest	7.9	6.3	Small airway epithelium TNFalpha + IL-1beta	9.9	4.8
CD45RA CD4 lymphocyte act	10.5	10.2	Coronary artery SMC rest	6.7	4.6
CD45RO CD4 lymphocyte act	0.0	17.8	Coronary artery SMC TNFalpha + IL-1beta	4.9	3.5
CD8 lymphocyte act	28.5	15.1	Astrocytes rest	11.6	5.2
Secondary CD8 lymphocyte rest	10.6	16.2	Astrocytes TNFalpha + IL-1beta	6.1	5.3
Secondary CD8 lymphocyte act	10.2	2.6	KU-812 (Basophil) rest	15.7	11.2
CD4 lymphocyte none	4.2	2.8	KU-812 (Basophil) PMA/ionomycin	18.9	16.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	5.7	4.7	CCD1106 (Keratinocytes) none	32.8	13.4
LAK cells rest	11.1	8.3	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	17.2	2.1
LAK cells IL-2	15.9	11.8	Liver cirrhosis	3.7	2.7
LAK cells IL-2+IL-12	14.1	11.2	NCI-H292 none	10.4	14.6
LAK cells IL-2+IFN gamma	18.4	11.9	NCI-H292 IL-4	14.9	19.6
LAK cells IL-2+ IL-18	23.0	8.4	NCI-H292 IL-9	18.2	27.2
LAK cells PMA/ionomycin	5.5	0.8	NCI-H292 IL-13	15.6	16.4
NK Cells IL-2 rest	11.7	6.1	NCI-H292 IFN gamma	26.2	17.0
Two Way MLR 3 day	13.5	7.4	HPAEC none	12.1	8.0
Two Way MLR 5 day	11.3	6.0	HPAEC TNF alpha + IL-1 beta	19.6	14.1
Two Way MLR 7 day	13.3	7.9	Lung fibroblast none	36.9	24.3
PBMC rest	2.8	1.0	Lung fibroblast TNF alpha + IL-1 beta	20.9	11.8
PBMC PWM	22.5	10.3	Lung fibroblast IL-4	37.4	17.6
PBMC PHA-L	21.2	9.0	Lung fibroblast IL-9	86.5	39.8
Ramos (B cell) none	54.7	23.8	Lung fibroblast IL-13	43.5	20.4
Ramos (B cell) ionomycin	53.2	25.9	Lung fibroblast IFN gamma	49.0	20.7
B lymphocytes PWM	24.1	17.4	Dermal fibroblast CCD1070 rest	46.3	15.2

B lymphocytes CD40L and IL-4	20.4	4.0	Dermal fibroblast CCD1070 TNF alpha	23.7	5.9
EOL-1 dbcAMP	21.8	2.9	Dermal fibroblast CCD1070 IL-1 beta	15.1	6.0
EOL-1 dbcAMP PMA/ionomycin	10.2	0.0	Dermal fibroblast IFN gamma	10.2	6.4
Dendritic cells none	7.5	4.0	Dermal fibroblast IL-4	31.4	10.7
Dendritic cells LPS	6.3	0.8	Dermal Fibroblasts rest	14.6	8.1
Dendritic cells anti- CD40	7.1	5.2	Neutrophils TNF α +LPS	10.4	1.5
Monocytes rest	3.8	4.2	Neutrophils rest	11.5	1.7
Monocytes LPS	4.7	0.1	Colon	3.5	2.5
Macrophages rest	11.7	9.9	Lung	9.9	3.7
Macrophages LPS	2.3	1.3	Thymus	11.2	18.8
HUVEC none	13.5	5.5	Kidney	100.0	100.0
HUVEC starved	11.5	8.2			

CNS_neurodegeneration_v1.0 Summary: Ag4927/Ag4928 These results confirm the expression of this gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. See Panel 1.5 for a discussion of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.5 Summary: Ag4927/Ag4928 Two experiments with two different probe and primer sets produce results that are in excellent agreement. Highest expression of this gene is detected in a lung cancer and a gastric cancer cell line (CTs=25-26). Moderate levels of expression of this gene is also seen in cluster of cancer cell lines derived from gastric, colon, lung, renal, breast, ovarian, prostate, squamous cell carcinoma, melanoma and brain cancers. Thus, expression of this gene could be used as a marker to detect the presence of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of gastric, colon, lung, renal, breast, ovarian, prostate, squamous cell carcinoma, melanoma and brain cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

In addition, this gene is expressed at moderate levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, therapeutic modulation of this gene product may be useful in the treatment of central nervous system disorders such as

5 Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4927/Ag4928 Highest expression of this gene is detected in kidney (CTs=28-29.5). This gene is expressed at moderate to low levels in a wide range of cell types of significance in the immune response in health and disease. These cells

10 include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement

15 with the expression profile in General_screening_panel_v1.5 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus

20 erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

D. CG137717-01: FLJ37712 is protein-like protein.

Expression of gene CG137717-01 was assessed using the primer-probe set Ag4929, described in Table DA. Results of the RTQ-PCR runs are shown in Tables DB, DC, DD and DE.

25 Table DA. Probe Name Ag4929

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctcttcacacctgcattccta-3'	22	1003	267
Probe	TET-5'- tcctctactttaccaaagtgaatactgg a-3'-TAMRA	30	1028	268
Reverse	5'-ccatggaatgtcatcaaaagag-3'	22	1059	269

Table DB. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag4929, Run 305464508	Tissue Name	Rel. Exp.(%) Ag4929, Run 305464508
I10967 COPD-F	57.0	I12427 Match Control Psoriasis-F	35.6
I10980 COPD-F	2.7	I12418 Psoriasis-M	71.2
I10968 COPD-M	73.7	I12723 Match Control Psoriasis-M	14.8
I10977 COPD-M	14.4	I12419 Psoriasis-M	82.4
I10989 Emphysema-F	46.3	I12424 Match Control Psoriasis-M	4.5
I10992 Emphysema-F	3.7	I12420 Psoriasis-M	34.4
I10993 Emphysema-F	28.9	I12425 Match Control Psoriasis-M	34.2
I10994 Emphysema-F	7.9	I04689 (MF) OA Bone- Backus	23.3
I10995 Emphysema-F	8.1	I04690 (MF) Adj "Normal" Bone-Backus	9.0
I10996 Emphysema-F	1.5	I04691 (MF) OA Synovium-Backus	6.0
I10997 Asthma-M	1.4	I04692 (BA) OA Cartilage-Backus	0.0
I11001 Asthma-F	4.2	I04694 (BA) OA Bone- Backus	5.9
I11002 Asthma-F	3.9	I04695 (BA) Adj "Normal" Bone-Backus	3.9
I11003 Atopic Asthma- F	11.0	I04696 (BA) OA Synovium-Backus	4.1
I11004 Atopic Asthma- F	13.0	I04700 (SS) OA Bone- Backus	16.3
I11005 Atopic Asthma- F	7.5	I04701 (SS) Adj "Normal" Bone-Backus	8.7
I11006 Atopic Asthma- F	0.5	I04702 (SS) OA Synovium-Backus	7.7
I11417 Allergy-M	4.6	I17093 OA Cartilage Rep7	15.4
I12347 Allergy-M	0.4	I12672 OA Bone5	30.6
I12349 Normal Lung-F	0.3	I12673 OA Synovium5	11.6
I12357 Normal Lung-F	4.2	I12674 OA Synovial Fluid cells5	23.3
I12354 Normal Lung- M	5.4	I17100 OA Cartilage Rep14	6.6
I12374 Crohns-F	34.9	I12756 OA Bone9	6.9
I12389 Match Control Crohns-F	53.6	I12757 OA Synovium9	14.4

112375 Crohns-F	33.4	112758 OA Synovial Fluid Cells9	7.8
112732 Match Control Crohns-F	40.1	117125 RA Cartilage Rep2	100.0
112725 Crohns-M	8.9	113492 Bone2 RA	10.7
112387 Match Control Crohns-M	26.6	113493 Synovium2 RA	9.1
112378 Crohns-M	1.2	113494 Syn Fluid Cells RA	7.8
112390 Match Control Crohns-M	36.1	113499 Cartilage4 RA	15.8
112726 Crohns-M	17.7	113500 Bone4 RA	24.7
112731 Match Control Crohns-M	43.8	113501 Synovium4 RA	7.0
112380 Ulcer Col-F	16.6	113502 Syn Fluid Cells4 RA	13.3
112734 Match Control Ulcer Col-F	88.3	113495 Cartilage3 RA	16.5
112384 Ulcer Col-F	54.7	113496 Bone3 RA	16.7
112737 Match Control Ulcer Col-F	7.1	113497 Synovium3 RA	8.9
112386 Ulcer Col-F	6.6	113498 Syn Fluid Cells3 RA	24.0
112738 Match Control Ulcer Col-F	3.3	117106 Normal Cartilage Rep20	5.6
112381 Ulcer Col-M	0.0	113663 Bone3 Normal	0.0
112735 Match Control Ulcer Col-M	5.7	113664 Synovium3 Normal	0.0
112382 Ulcer Col-M	94.0	113665 Syn Fluid Cells3 Normal	0.5
112394 Match Control Ulcer Col-M	4.1	117107 Normal Cartilage Rep22	10.9
112383 Ulcer Col-M	36.6	113667 Bone4 Normal	14.8
112736 Match Control Ulcer Col-M	37.4	113668 Synovium4 Normal	12.0
112423 Psoriasis-F	26.6	113669 Syn Fluid Cells4 Normal	17.1

Table DC. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4929, Run 224735010	Tissue Name	Rel. Exp.(%) Ag4929, Run 224735010
AD I Hippo	5.9	Control (Path) 3 Temporal Ctx	1.1

AD 2 Hippo	8.0	Control (Path) 4 Temporal Ctx	25.7
AD 3 Hippo	2.2	AD 1 Occipital Ctx	10.8
AD 4 Hippo	1.0	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	100.0	AD 3 Occipital Ctx	2.6
AD 6 Hippo	26.6	AD 4 Occipital Ctx	10.7
Control 2 Hippo	8.8	AD 5 Occipital Ctx	17.6
Control 4 Hippo	1.2	AD 6 Occipital Ctx	18.7
Control (Path) 3 Hippo	1.0	Control 1 Occipital Ctx	0.3
AD 1 Temporal Ctx	6.3	Control 2 Occipital Ctx	48.0
AD 2 Temporal Ctx	21.2	Control 3 Occipital Ctx	14.2
AD 3 Temporal Ctx	1.6	Control 4 Occipital Ctx	1.0
AD 4 Temporal Ctx	9.5	Control (Path) 1 Occipital Ctx	59.5
AD 5 Inf Temporal Ctx	90.1	Control (Path) 2 Occipital Ctx	9.9
AD 5 Sup Temporal Ctx	27.7	Control (Path) 3 Occipital Ctx	0.4
AD 6 Inf Temporal Ctx	43.5	Control (Path) 4 Occipital Ctx	17.1
AD 6 Sup Temporal Ctx	41.8	Control 1 Parietal Ctx	1.9
Control 1 Temporal Ctx	1.5	Control 2 Parietal Ctx	33.2
Control 2 Temporal Ctx	24.0	Control 3 Parietal Ctx	17.3
Control 3 Temporal Ctx	10.4	Control (Path) 1 Parietal Ctx	42.6
Control 4 Temporal Ctx	2.3	Control (Path) 2 Parietal Ctx	16.3
Control (Path) 1 Temporal Ctx	39.8	Control (Path) 3 Parietal Ctx	0.9
Control (Path) 2 Temporal Ctx	27.7	Control (Path) 4 Parietal Ctx	31.6

Table DD. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag4929, Run 228839297	Tissue Name	Rel. Exp.(%) Ag4929, Run 228839297
Adipose	0.0	Renal ca. TK-10	65.5
Melanoma* Hs688(A).T	0.0	Bladder	0.1
Melanoma* Hs688(B).T	0.1	Gastric ca. (liver met.) NCI-N87	3.3
Melanoma* M14	40.9	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	7.2

Melanoma* SK-MEL-5	0.0	Colon ca. SW480	1.8
Squamous cell carcinoma SCC-4	17.1	Colon ca.* (SW480 met) SW620	0.2
Testis Pool	1.4	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	40.9	Colon ca. HCT-116	46.7
Prostate Pool	1.0	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.2
Uterus Pool	0.7	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	1.7	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	88.3	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.5	Colon Pool	0.9
Ovarian ca. OVCAR-5	44.1	Small Intestine Pool	1.4
Ovarian ca. IGROV-1	3.4	Stomach Pool	1.5
Ovarian ca. OVCAR-8	9.1	Bone Marrow Pool	0.0
Ovary	3.3	Fetal Heart	0.0
Breast ca. MCF-7	0.3	Heart Pool	1.6
Breast ca. MDA-MB-231	67.8	Lymph Node Pool	3.0
Breast ca. BT 549	0.4	Fetal Skeletal Muscle	0.2
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.6
Breast ca. MDA-N	0.2	Spleen Pool	1.1
Breast Pool	1.6	Thymus Pool	3.5
Trachea	3.5	CNS cancer (glio/astro) U87-MG	0.2
Lung	0.0	CNS cancer (glio/astro) U-118-MG	1.2
Fetal Lung	2.5	CNS cancer (neuro;met) SK-N-AS	48.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	2.1
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	5.6
Lung ca. SHP-77	6.1	CNS cancer (glio) SF-295	0.4
Lung ca. A549	32.8	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	5.3	Brain (cerebellum)	14.5
Lung ca. NCI-H23	1.3	Brain (fetal)	4.7
Lung ca. NCI-H460	21.2	Brain (Hippocampus) Pool	10.1
Lung ca. HOP-62	15.0	Cerebral Cortex Pool	0.5
Lung ca. NCI-H522	0.6	Brain (Substantia nigra) Pool	1.7
Liver	0.0	Brain (Thalamus) Pool	27.9
Fetal Liver	2.9	Brain (whole)	39.0

Liver ca. HepG2	0.0	Spinal Cord Pool	14.7
Kidney Pool	8.3	Adrenal Gland	1.7
Fetal Kidney	2.2	Pituitary gland Pool	4.0
Renal ca. 786-0	3.3	Salivary Gland	5.9
Renal ca. A498	12.0	Thyroid (female)	0.9
Renal ca. ACHN	51.1	Pancreatic ca. CAPAN2	100.0
Renal ca. UO-31	19.8	Pancreas Pool	1.6

Table DE. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4929, Run 223597249	Tissue Name	Rel. Exp.(%) Ag4929, Run 223597249
Secondary Th1 act	55.1	HUVEC IL-1beta	0.0
Secondary Th2 act	22.2	HUVEC IFN gamma	0.1
Secondary Tr1 act	34.4	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	3.6	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	3.8	HUVEC IL-11	0.7
Secondary Tr1 rest	13.8	Lung Microvascular EC none	0.4
Primary Th1 act	60.3	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	40.1	Microvascular Dermal EC none	0.0
Primary Tr1 act	100.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	10.4	Bronchial epithelium TNFalpha + IL1beta	4.5
Primary Th2 rest	9.6	Small airway epithelium none	5.4
Primary Tr1 rest	42.6	Small airway epithelium TNFalpha + IL-1beta	7.2
CD45RA CD4 lymphocyte act	7.2	Coronary artery SMC rest	0.4
CD45RO CD4 lymphocyte act	5.9	Coronary artery SMC TNFalpha + IL-1beta	2.2
CD8 lymphocyte act	11.3	Astrocytes rest	3.4
Secondary CD8 lymphocyte rest	5.7	Astrocytes TNFalpha + IL- 1beta	2.6
Secondary CD8 lymphocyte act	14.7	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	2.0	KU-812 (Basophil) PMA/ionomycin	0.2
2ry Th1/Th2/Tr1_anti- CD95 CH11	9.7	CCD1106 (Keratinocytes) none	14.1

LAK cells rest	4.1	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	6.7
LAK cells IL-2	1.0	Liver cirrhosis	0.4
LAK cells IL-2+IL-12	6.9	NCI-H292 none	3.2
LAK cells IL-2+IFN gamma	12.7	NCI-H292 IL-4	2.5
LAK cells IL-2+ IL-18	16.2	NCI-H292 IL-9	4.4
LAK cells PMA/ionomycin	0.3	NCI-H292 IL-13	2.9
NK Cells IL-2 rest	4.1	NCI-H292 IFN gamma	1.7
Two Way MLR 3 day	3.6	HPAEC none	0.0
Two Way MLR 5 day	8.1	HPAEC TNF alpha + IL-1 beta	0.2
Two Way MLR 7 day	3.4	Lung fibroblast none	0.3
PBMC rest	0.7	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.7	Lung fibroblast IL-4	0.0
PBMC PHA-L	5.8	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	16.4	Dermal fibroblast CCD1070 rest	0.2
B lymphocytes CD40L and IL-4	0.9	Dermal fibroblast CCD1070 TNF alpha	3.5
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	1.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.5
Dendritic cells none	1.4	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.4	Dermal Fibroblasts rest	0.2
Dendritic cells anti-CD40	1.0	Neutrophils TNFa+LPS	0.5
Monocytes rest	1.0	Neutrophils rest	0.9
Monocytes LPS	0.2	Colon	0.3
Macrophages rest	0.8	Lung	2.0
Macrophages LPS	0.2	Thymus	8.1
HUVEC none	0.0	Kidney	42.3
HUVEC starved	0.0		

AI_comprehensive_panel_v1.0 Summary: Ag4929 Highest expression of this gene is detected in RA cartilage (CT=30.6). In addition, moderate levels of expression are seen in samples from Crohn's, ulcerative colitis, psoriasis, and COPD derived tissue. Thus, modulation of the expression or function of this gene may be useful in the treatment of these

5 conditions.

CNS_neurodegeneration_v1.0 Summary: Ag4929 This panel does not show differential expression of this gene in Alzheimer's disease. However, this profile confirms the expression of this gene at moderate levels in the brain. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurological disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

General_screening_panel_v1.5 Summary: Ag4929 Highest expression of this gene is seen in a pancreatic cancer cell line (CT=28). Expression in this panel appears to be predominantly associated with samples derived from cancer cell lines, including brain, renal, lung, breast, ovarian, prostate and melanoma cancer cell lines. Thus, expression of this gene could be used as a marker of cancer. Furthermore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of cancer.

Panel 4.1D Summary: Ag4929 Expression of this gene is most prominent in T cells including both acutely and chronically activated T cells (CTs=29-30). Therefore, therapeutics designed with the protein encoded by this transcript may help to regulate T cell function and be effective in treating T cell mediated diseases such as asthma, arthritis, psoriasis, , and lupus.

E. CG137793-02: High Affinity Immunoglobulin Epsilon

Receptor Alpha-Subunit Precursor Protein-like Protein.

Expression of gene CG137793-02 was assessed using the primer-probe set Ag6866, described in Table EA.

Table EA. Probe Name Ag6866

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agaatacaaatgccatgggtt-3'	20	292	270
Probe	TET-5'- tcctataatagatcaccttgtagacatcc ca-3'-TAMRA	32	320	271
Reverse	5'-ggttctcataccagtacttgaga-3'	23	361	272

F. CG137873-02: Human fibrinogen alpha chain precursor protein-likew protein

Expression of gene CG137873-02 was assessed using the primer-probe set Ag7411, described in Table FA. Results of the RTQ-PCR runs are shown in Table FB.

Table FA. Probe Name Ag7411

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - acccagactggggctca - 3'	17	1196	273
Probe	TET- 5' - atctggcatcttcacaaatacaaagg - 3' - TAMRA	26	1215	274
Reverse	5' - atttaccacgggaagggaa - 3'	19	1273	275

5

Table FB. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag7411, Run 305065220	Tissue Name	Rel. Exp.(%) Ag7411, Run 305065220
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL- 1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0

CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti- CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	100.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	0.0	Colon	0.0
Macrophages rest	0.0	Lung	0.0
Macrophages LPS	0.0	Thymus	0.0
HUVEC none	0.0	Kidney	0.0
HUVEC starved	0.0		

Panel 4.1D Summary: Ag7411 Significant expression of this gene is detected in a liver cirrhosis sample (CT = 33.8). Furthermore, expression of this gene is not detected in normal liver on Panel 1.6, suggesting that its expression is unique to liver cirrhosis.

Therefore, therapeutic modulation of the expression or function of this gene may be used to diagnose this condition or to reduce or inhibit fibrosis that occurs in liver cirrhosis.

G. CG137873-03 (205101513edited2): Fibrinogen Alpha Chain

Precursor Protein-like Protein.

- 5 Expression of gene CG137873-03 (205101513edited2) was assessed using the primer-probe set Ag7412, described in Table GA. Results of the RTQ-PCR runs are shown in Tables GB and GC.

Table GA. Probe Name Ag7412

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ggaagctggaagctggaagta-3'	21	970	276
Probe	TET-5'- ccaaaaccctgggagccctagacctg -3'-TAMRA	26	998	277
Reverse	5'-ctgccaggattccagggtt-3'	18	1034	278

Table GB. General_screening_panel_v1.6

Tissue Name	Rel. Exp.(%) Ag7412, Run 306067375	Tissue Name	Rel. Exp.(%) Ag7412, Run 306067375
Adipose	0.0	Renal ca. TK-10	3.3
Melanoma* Hs688(A).T	0.0	Bladder	1.0
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	0.0	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.0
Prostate Pool	0.0	Colon ca. CaCo-2	6.4
Placenta	0.0	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.0
Ovarian ca. OVCAR-5	0.0	Small Intestine Pool	0.0
Ovarian ca. IGROV-1	0.0	Stomach Pool	0.1

Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.0
Ovary	0.0	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.0
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.0
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	0.0
Breast Pool	0.0	Thymus Pool	0.0
Trachea	0.0	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	0.5	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF- 539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB- 75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB- 19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	4.6	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	0.0	Brain (fetal)	0.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	0.0
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	0.0
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.0
Liver	0.0	Brain (Thalamus) Pool	0.0
Fetal Liver	100.0	Brain (whole)	4.6
Liver ca. HepG2	6.7	Spinal Cord Pool	0.0
Kidney Pool	0.0	Adrenal Gland	0.0
Fetal Kidney	0.0	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.0	Pancreas Pool	0.0

Table GC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag7412, Run 305065272	Tissue Name	Rel. Exp.(%) Ag7412, Run 305065272
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0

Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	100.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0

PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	0.0	Colon	0.0
Macrophages rest	0.0	Lung	0.0
Macrophages LPS	0.0	Thymus	0.0
HUVEC none	0.0	Kidney	0.0
HUVEC starved	0.0		

General_screening_panel_v1.6 Summary: Ag7412 Highest expression of this gene is seen in fetal liver (CT=27). Thus, expression of this gene could be used to differentiate between fetal and adult liver (CT=40). Furthermore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of liver disorders.

- 5 **Panel 4.1D Summary:** Ag7412 Significant expression of this gene is detected in a liver cirrhosis sample (CT = 28.3). Therefore, therapeutic modulation of the expression or function of this gene may be used to diagnose this condition and to reduce or inhibit fibrosis that occurs in liver cirrhosis.

H. CG137882-02: Membrane Protein FLJ212269-like Protein.

- 10 Expression of gene CG137882-02 was assessed using the primer-probe set Ag7046, described in Table HA.

Table HA. Probe Name Ag7046

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - tgtgaacgtcgaagcaacc - 3'	19	391	279
Probe	TET- 5' - agtctcaccttcacgcgacaagcttc - 3' - TAMRA	27	421	280

Reverse	5' - tgggagagatatattggaaaggaat - 3'	23	461	281
---------	--	----	-----	-----

General_screening_panel_v1.6 Summary: Ag7046 Expression of this gene is low/undetectable in all samples on this panel (CTs>35).

I. CG137910-01: FLJ21432-like protein.

Expression of gene CG137910-01 was assessed using the primer-probe set Ag7448,
5 described in Table IA. Results of the RTQ-PCR runs are shown in Tables IB and IC.

Table IA. Probe Name Ag7448

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - aggagccattctctgccttt - 3'	20	315	282
Probe	TET-5' - catggctcttccacacagtctactgcca -3' - TAMRA	28	341	283
Reverse	5' - cagtttagagaagagccgagaga - 3'	23	380	284

Table IB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag7448, Run 306067416	Tissue Name	Rel. Exp.(%) Ag7448, Run 306067416
AD 1 Hippo	12.5	Control (Path) 3 Temporal Ctx	5.2
AD 2 Hippo	29.1	Control (Path) 4 Temporal Ctx	12.2
AD 3 Hippo	9.9	AD 1 Occipital Ctx	14.3
AD 4 Hippo	6.4	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	48.3	AD 3 Occipital Ctx	16.3
AD 6 Hippo	41.2	AD 4 Occipital Ctx	17.3
Control 2 Hippo	16.4	AD 5 Occipital Ctx	11.2
Control 4 Hippo	10.4	AD 6 Occipital Ctx	25.5
Control (Path) 3 Hippo	2.9	Control 1 Occipital Ctx	5.8
AD 1 Temporal Ctx	7.3	Control 2 Occipital Ctx	30.8
AD 2 Temporal Ctx	27.5	Control 3 Occipital Ctx	11.6
AD 3 Temporal Ctx	9.5	Control 4 Occipital Ctx	11.6
AD 4 Temporal Ctx	16.2	Control (Path) 1 Occipital Ctx	47.0
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	5.1
AD 5 SupTemporal Ctx	52.9	Control (Path) 3 Occipital Ctx	9.5

AD 6 Inf Temporal Ctx	58.2	Control (Path) 4 Occipital Ctx	9.6
AD 6 Sup Temporal Ctx	36.1	Control 1 Parietal Ctx	4.1
Control 1 Temporal Ctx	5.6	Control 2 Parietal Ctx	37.9
Control 2 Temporal Ctx	32.1	Control 3 Parietal Ctx	10.2
Control 3 Temporal Ctx	8.5	Control (Path) 1 Parietal Ctx	27.9
Control 4 Temporal Ctx	7.4	Control (Path) 2 Parietal Ctx	12.8
Control (Path) 1 Temporal Ctx	20.6	Control (Path) 3 Parietal Ctx	9.8
Control (Path) 2 Temporal Ctx	16.6	Control (Path) 4 Parietal Ctx	17.2

Table IC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag7448, Run 306067435	Tissue Name	Rel. Exp.(%) Ag7448, Run 306067435
Secondary Th1 act	53.2	HUVEC IL-1beta	55.1
Secondary Th2 act	81.2	HUVEC IFN gamma	50.7
Secondary Tr1 act	8.6	HUVEC TNF alpha + IFN gamma	8.2
Secondary Th1 rest	5.6	HUVEC TNF alpha + IL4	21.6
Secondary Th2 rest	6.7	HUVEC IL-11	20.3
Secondary Tr1 rest	3.8	Lung Microvascular EC none	84.1
Primary Th1 act	10.7	Lung Microvascular EC TNFalpha + IL-1beta	24.3
Primary Th2 act	51.8	Microvascular Dermal EC none	18.0
Primary Tr1 act	62.9	Microvascular Dermal EC TNFalpha + IL-1beta	11.0
Primary Th1 rest	3.7	Bronchial epithelium TNFalpha + IL1beta	26.1
Primary Th2 rest	5.0	Small airway epithelium none	29.9
Primary Tr1 rest	0.6	Small airway epithelium TNFalpha + IL-1beta	58.6
CD45RA CD4 lymphocyte act	45.7	Coronary artery SMC rest	16.0
CD45RO CD4 lymphocyte act	68.3	Coronary artery SMC TNFalpha + IL-1beta	23.8
CD8 lymphocyte act	10.4	Astrocytes rest	5.5
Secondary CD8 lymphocyte rest	12.3	Astrocytes TNFalpha + IL-1beta	4.8
Secondary CD8 lymphocyte act	12.0	KU-812 (Basophil) rest	34.2

CD4 lymphocyte none	1.0	KU-812 (Basophil) PMA/ionomycin	79.0
2ry Th1/Th2/Tr1_anti- CD95 CH11	9.8	CCD1106 (Keratinocytes) none	16.8
LAK cells rest	18.2	CCD1106 (Keratinocytes) TNFalpha + IL-1 beta	6.9
LAK cells IL-2	18.9	Liver cirrhosis	9.4
LAK cells IL-2+IL-12	1.3	NCI-H292 none	40.3
LAK cells IL-2+IFN gamma	9.2	NCI-H292 IL-4	68.8
LAK cells IL-2+ IL-18	10.4	NCI-H292 IL-9	75.3
LAK cells PMA/ionomycin	77.9	NCI-H292 IL-13	39.5
NK Cells IL-2 rest	82.4	NCI-H292 IFN gamma	19.6
Two Way MLR 3 day	9.5	HPAEC none	7.2
Two Way MLR 5 day	5.4	HPAEC TNF alpha + IL-1 beta	58.6
Two Way MLR 7 day	10.4	Lung fibroblast none	36.3
PBMC rest	3.8	Lung fibroblast TNF alpha + IL-1 beta	23.7
PBMC PWM	20.9	Lung fibroblast IL-4	30.8
PBMC PHA-L	15.6	Lung fibroblast IL-9	55.5
Ramos (B cell) none	33.4	Lung fibroblast IL-13	7.6
Ramos (B cell) ionomycin	97.3	Lung fibroblast IFN gamma	51.4
B lymphocytes PWM	32.5	Dermal fibroblast CCD1070 rest	63.7
B lymphocytes CD40L and IL-4	45.4	Dermal fibroblast CCD1070 TNF alpha	100.0
EOL-1 dbcAMP	64.2	Dermal fibroblast CCD1070 IL-1 beta	40.9
EOL-1 dbcAMP PMA/ionomycin	13.1	Dermal fibroblast IFN gamma	24.7
Dendritic cells none	10.7	Dermal fibroblast IL-4	30.8
Dendritic cells LPS	4.9	Dermal Fibroblasts rest	23.2
Dendritic cells anti-CD40	18.2	Neutrophils TNFa+LPS	7.6
Monocytes rest	7.4	Neutrophils rest	15.4
Monocytes LPS	79.0	Colon	8.1
Macrophages rest	17.8	Lung	4.8
Macrophages LPS	13.2	Thymus	4.7
HUVEC none	27.9	Kidney	18.6
HUVEC starved	51.1		

CNS_neurodegeneration_v1.0 Summary: Ag7448 This gene appears to be upregulated in the temporal cortex of Alzheimer's disease patients when compared with

non-demented controls. Therefore, modulation of the expression or function of this gene may slow or stop the progression of Alzheimer's disease.

Panel 4.1D Summary: Ag7448 This gene is ubiquitously expressed in this panel with highest expression in TNF- α treated dermal fibroblasts (CT=29). This gene is also expressed at moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues as well as in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

J. CG138013-01: Sialic acid-binding immunoglobulin like lectin-9-like protein.

Expression of gene CG138013-01 was assessed using the primer-probe set Ag4957, described in Table JA. Results of the RTQ-PCR runs are shown in Tables JB, JC and JD.

Table JA. Probe Name Ag4957

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-cggggagatacttctttcgtat-3'	22	422	285
Probe	TET-5'- tgaattataaacatcacggctctctg -3'-TAMRA	28	463	286
Reverse	5'-ggtcaaggctgtcacattca-3'	20	491	287

Table JB. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag4957, Run 222176655	Tissue Name	Rel. Exp.(%) Ag4957, Run 222176655
110967 COPD-F	6.1	112427 Match Control Psoriasis-F	5.7
110980 COPD-F	3.2	112418 Psoriasis-M	4.4

110968 COPD-M	6.4	112723 Match Control Psoriasis-M	2.6
110977 COPD-M	4.0	112419 Psoriasis-M	9.0
110989 Emphysema-F	4.2	112424 Match Control Psoriasis-M	4.2
110992 Emphysema-F	2.3	112420 Psoriasis-M	8.1
110993 Emphysema-F	3.8	112425 Match Control Psoriasis-M	3.5
110994 Emphysema-F	1.4	104689 (MF) OA Bone-Backus	29.3
110995 Emphysema-F	7.1	104690 (MF) Adj "Normal" Bone-Backus	11.1
110996 Emphysema-F	2.7	104691 (MF) OA Synovium-Backus	37.9
110997 Asthma-M	5.1	104692 (BA) OA Cartilage-Backus	0.9
111001 Asthma-F	6.7	104694 (BA) OA Bone-Backus	29.5
111002 Asthma-F	5.6	104695 (BA) Adj "Normal" Bone-Backus	10.4
111003 Atopic Asthma-F	1.4	104696 (BA) OA Synovium-Backus	100.0
111004 Atopic Asthma-F	4.0	104700 (SS) OA Bone-Backus	31.0
111005 Atopic Asthma-F	2.0	104701 (SS) Adj "Normal" Bone-Backus	19.2
111006 Atopic Asthma-F	0.5	104702 (SS) OA Synovium-Backus	28.3
111417 Allergy-M	4.0	117093 OA Cartilage Rep7	4.0
112347 Allergy-M	0.0	112672 OA Bone5	8.8
112349 Normal Lung-F	0.5	112673 OA Synovium5	4.5
112357 Normal Lung-F	16.7	112674 OA Synovial Fluid cells5	2.7
112354 Normal Lung-M	7.7	117100 OA Cartilage Rep14	3.1
112374 Crohns-F	7.2	112756 OA Bone9	6.0
112389 Match Control Crohns-F	5.5	112757 OA Synovium9	0.7
112375 Crohns-F	1.6	112758 OA Synovial Fluid Cells9	5.5
112732 Match Control Crohns-F	5.2	117125 RA Cartilage Rep2	5.0
112725 Crohns-M	1.1	113492 Bone2 RA	17.9
112387 Match Control Crohns-M	3.4	113493 Synovium2 RA	8.1

112378 Crohns-M	0.8	113494 Syn Fluid Cells RA	14.0
112390 Match Control Crohns-M	3.9	113499 Cartilage4 RA	17.4
112726 Crohns-M	2.2	113500 Bone4 RA	16.3
112731 Match Control Crohns-M	1.7	113501 Synovium4 RA	12.5
112380 Ulcer Col-F	1.9	113502 Syn Fluid Cells4 RA	10.7
112734 Match Control Ulcer Col-F	15.5	113495 Cartilage3 RA	24.7
112384 Ulcer Col-F	6.7	113496 Bone3 RA	23.8
112737 Match Control Ulcer Col-F	1.2	113497 Synovium3 RA	12.6
112386 Ulcer Col-F	1.1	113498 Syn Fluid Cells3 RA	26.2
112738 Match Control Ulcer Col-F	4.1	117106 Normal Cartilage Rep20	2.5
112381 Ulcer Col-M	0.0	113663 Bone3 Normal	0.7
112735 Match Control Ulcer Col-M	5.6	113664 Synovium3 Normal	0.0
112382 Ulcer Col-M	5.1	113665 Syn Fluid Cells3 Normal	0.0
112394 Match Control Ulcer Col-M	0.4	117107 Normal Cartilage Rep22	1.7
112383 Ulcer Col-M	15.1	113667 Bone4 Normal	1.8
112736 Match Control Ulcer Col-M	1.4	113668 Synovium4 Normal	1.5
112423 Psoriasis-F	4.0	113669 Syn Fluid Cells4 Normal	4.1

Table JC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4957, Run 219311035	Tissue Name	Rel. Exp.(%) Ag4957, Run 219311035
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.1	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.2
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0

Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.1	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.1
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.4	Coronary artery SMC TNFalpha + IL-1beta	0.2
CD8 lymphocyte act	0.3	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.4	Astrocytes TNFalpha + IL- 1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.5	KU-812 (Basophil) PMA/ionomycin	0.4
2ry Th1/Th2/Tr1_anti- CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	17.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.2	Liver cirrhosis	1.0
LAK cells IL-2+IL-12	0.4	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.5	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.3	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	8.5	NCI-H292 IL-13	0.1
NK Cells IL-2 rest	1.7	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	10.4	HPAEC none	0.0
Two Way MLR 5 day	3.9	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.8	Lung fibroblast none	0.0
PBMC rest	7.5	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	2.4	Lung fibroblast IL-4	0.0
PBMC PHA-L	4.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.2	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.1	Dermal fibroblast CCD1070 TNF alpha	0.1
EOL-1 dbcAMP	1.4	Dermal fibroblast CCD1070 IL-1 beta	0.0

EOL-1 dbcAMP PMA/ionomycin	5.7	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	35.6	Dermal fibroblast IL-4	0.3
Dendritic cells LPS	25.5	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	57.4	Neutrophils TNFa+LPS	3.9
Monocytes rest	51.8	Neutrophils rest	40.3
Monocytes LPS	100.0	Colon	0.1
Macrophages rest	29.1	Lung	19.9
Macrophages LPS	12.2	Thymus	0.5
HUVEC none	0.0	Kidney	0.1
HUVEC starved	0.0		

Table JD. general oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4957, Run 260281958	Tissue Name	Rel. Exp.(%) Ag4957, Run 260281958
Colon cancer 1	20.9	Bladder cancer NAT 2	0.0
Colon cancer NAT 1	14.6	Bladder cancer NAT 3	0.0
Colon cancer 2	34.9	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	5.5	Prostate adenocarcinoma 1	13.5
Colon cancer 3	22.5	Prostate adenocarcinoma 2	1.5
Colon cancer NAT 3	4.2	Prostate adenocarcinoma 3	3.3
Colon malignant cancer 4	44.1	Prostate adenocarcinoma 4	17.1
Colon normal adjacent tissue 4	5.6	Prostate cancer NAT 5	3.0
Lung cancer 1	33.4	Prostate adenocarcinoma 6	1.4
Lung NAT 1	12.2	Prostate adenocarcinoma 7	3.0
Lung cancer 2	22.4	Prostate adenocarcinoma 8	0.0
Lung NAT 2	4.8	Prostate adenocarcinoma 9	10.4
Squamous cell carcinoma 3	43.2	Prostate cancer NAT 10	0.0
Lung NAT 3	2.9	Kidney cancer 1	73.2
metastatic melanoma 1	16.6	Kidney NAT 1	11.1
Melanoma 2	3.2	Kidney cancer 2	80.7
Melanoma 3	0.0	Kidney NAT 2	4.0
metastatic melanoma 4	70.7	Kidney cancer 3	20.0
metastatic melanoma 5	100.0	Kidney NAT 3	3.1
Bladder cancer 1	3.5	Kidney cancer 4	23.5

Bladder cancer NAT 1	0.0	Kidney NAT 4	5.0
Bladder cancer 2	4.2		

AI_comprehensive panel_v1.0 Summary: Ag4957 Highest expression of this gene is detected in orthoarthritis synovium (CT=31.5). In addition, moderate to low levels of expression of this gene is also seen in samples derived from osteoarthritic (OA) bone and adjacent bone as well as OA and normal bone, and OA synovium. Low level expression is also detected in cartilage, bone, synovium and synovial fluid samples from rheumatoid arthritis patients. This gene codes for a variant of sialic acid-binding immunoglobulin-like lectin-9 (SIGLEC-9) protein. Siglec-9 was found to be expressed at high or intermediate levels by monocytes, neutrophils, and a minor population of CD16(+), CD56(-) cells and at lower levels in B cells, NK cells and minor subsets of CD8(+) T cells and CD4(+) T cells (Zhang *et al.*, 2000, J Biol Chem 275(29):22121-6, PMID: 10801862). Similar pattern of expression of SIGLEC-9 encoded by this gene in monocytes, neutrophils and T cells, is also seen in panel 4.1D. Monocytes and T cells are known to play a role in the pathogenesis of arthritis (VanderBorghet *et al.*, 2001, Semin Arthritis Rheum 31(3):160-75, PMID: 11740797; Jenkins JK *et al.*, 2002, Am J Med Sci 323(4):171-80, PMID: 12003371). Therefore, therapeutic modulation of the SIGLEC-9 protein encoded by this gene may be useful in the treatment of osteoarthritis and rheumatoid arthritis.

Panel 4.1D Summary: Ag4957 Highest expression of this gene is detected in LPS treated monocytes (CT=28.5). In addition, moderate levels of expression of this gene is also seen in resting monocytes, dendritic cell, and macrophages. Thus, therapeutic modalities that block the function of the this gene product may be useful in the reduction or elimination of the symptoms in patients with autoimmune and inflammatory diseases in which monocytes, dendritic cells and macrophages play an important role in antigen presentation and other functions. Furthermore, moderate to low levels of expression of this gene is also seen in eosinophils, PBMC cells, two way MLR, LAK cells, stimulated neutrophils and lung. Therefore, therapeutic modulation of this gene product may be beneficial in the treatment of autoimmune and inflammatory diseases, such as lupus erythematosus, Crohn's disease, ulcerative colitis, multiple sclerosis, chronic obstructive pulmonary disease, asthma, emphysema, or rheumatoid arthritis.

general oncology screening panel_v_2.4 Summary: Ag4957 Highest expression of this gene is detected in metastatic melanoma (CT=33.2). Moderate to low levels of

expression of this gene is also seen in malignant colon cancer, lung cancer, and kidney cancer. Expression of this gene is higher in cancer as compared to the corresponding adjacent normal tissue. Therefore, expression of this gene may be used as diagnostic marker for detection of these cancers and therapeutic modulation of this gene or its product through the use of small molecule drug or antibodies may be useful in the treatment of these cancers and also their metastasis.

K. CG138074-01: RIKEN 2310012P03-like protein.

Expression of gene CG138074-01 was assessed using the primer-probe set Ag4952, described in Table KA.

10 Table KA. Probe Name Ag4952

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tacaccaccatgctgtccat-3'	20	574	288
Probe	TET-5'- ccatatccattctgccttggacacct -3'-TAMRA	26	609	289
Reverse	5'-actcgtgtcactcatcatgtca- 3'	22	648	290

L. CG138573-01: FOLATE RECEPTOR 3-LIKE PROTEIN.

Expression of gene CG138573-01 was assessed using the primer-probe set Ag4964, described in Table LA.

15 Table LA. Probe Name Ag4964

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctggatgtatccccactctaca- 3'	22	256	291
Probe	TET-5'- ttcagcctgtttcactgtggactgct -3'-TAMRA	26	280	292
Reverse	5'-tagaagcagatagcctggatga- 3'	22	329	293

General_screening_panel_v1.5 Summary: Ag4964 Expression of this gene is low/undetectable in all samples on this panel (CT>35).

Panel 4.1D Summary: Ag4964 Expression of this gene is low/undetectable in all samples on this panel (CTs>35).

**M. CG138606-01: BRUSH BORDER 61.9 KDA PROTEIN
PRECURSOR-LIKE PROTEIN.**

Expression of gene CG138606-01 was assessed using the primer-probe set Ag4970, described in Table MA. Results of the RTQ-PCR runs are shown in Tables MB and MC.

5 Table MA. Probe Name Ag4970

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' -ttatatccctcgggaaattgac-3'	22	1248	294
Probe	TET-5' - aaacacagccatcgatcacctttg -3' -TAMRA	26	1287	295
Reverse	5' -tgtcaatgggaaatggtctaaa-3'	22	1321	296

Table MB. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag4970, Run 228926385	Tissue Name	Rel. Exp.(%) Ag4970, Run 228926385
Adipose	5.1	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	1.4	Bladder	2.9
Melanoma* Hs688(B).T	3.1	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	4.1
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	1.3
Squamous cell carcinoma SCC-4	0.7	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	21.6	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.7	Colon ca. HCT-116	1.2
Prostate Pool	2.7	Colon ca. CaCo-2	2.6
Placenta	0.0	Colon cancer tissue	3.5
Uterus Pool	3.4	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	10.7	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.6	Colon Pool	16.7
Ovarian ca. OVCAR-5	0.0	Small Intestine Pool	100.0
Ovarian ca. IGROV-1	0.5	Stomach Pool	4.9
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	6.0
Ovary	5.0	Fetal Heart	4.2
Breast ca. MCF-7	0.5	Heart Pool	3.2
Breast ca. MDA-MB-231	0.5	Lymph Node Pool	13.5

Breast ca. BT 549	5.4	Fetal Skeletal Muscle	1.8
Breast ca. T47D	0.0	Skeletal Muscle Pool	5.7
Breast ca. MDA-N	0.0	Spleen Pool	3.3
Breast Pool	9.4	Thymus Pool	8.0
Trachea	0.9	CNS cancer (glio/astro) U87-MG	0.0
Lung	4.3	CNS cancer (glio/astro) U-118-MG	2.2
Fetal Lung	5.1	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF- 539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB- 75	4.4
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB- 19	0.5
Lung ca. SHP-77	1.1	CNS cancer (glio) SF-295	1.7
Lung ca. A549	0.0	Brain (Amygdala) Pool	1.5
Lung ca. NCI-H526	0.0	Brain (cerebellum)	1.2
Lung ca. NCI-H23	1.2	Brain (fetal)	1.3
Lung ca. NCI-H460	0.8	Brain (Hippocampus) Pool	2.0
Lung ca. HOP-62	0.9	Cerebral Cortex Pool	1.9
Lung ca. NCI-H522	0.3	Brain (Substantia nigra) Pool	1.1
Liver	0.0	Brain (Thalamus) Pool	0.9
Fetal Liver	3.7	Brain (whole)	0.5
Liver ca. HepG2	0.6	Spinal Cord Pool	2.1
Kidney Pool	21.0	Adrenal Gland	0.4
Fetal Kidney	13.6	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	4.6
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	1.1	Pancreas Pool	9.5

Table MC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4970, Run 223692673	Tissue Name	Rel. Exp.(%) Ag4970, Run 223692673
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.8
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	3.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0

Secondary Tr1 rest	1.1	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1 beta	0.9
Primary Th2 act	0.0	Microvascular Dermal EC none	0.4
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1 beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1 beta	2.6
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	1.1	Small airway epithelium TNFalpha + IL-1 beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1 beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1 beta	0.7
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	3.2
CD4 lymphocyte none	0.8	KU-812 (Basophil) PMA/ionomycin	8.4
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	1.2
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1 beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	100.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.9
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	1.3
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	1.8
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.8
Two Way MLR 7 day	0.0	Lung fibroblast none	3.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	1.5
PBMC PWM	0.0	Lung fibroblast IL-4	0.9
PBMC PHA-L	0.0	Lung fibroblast IL-9	2.8
Ramos (B cell) none	0.0	Lung fibroblast IL-13	1.8
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.9

B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.8
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	3.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	1.7
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	2.4
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	0.0	Colon	15.6
Macrophages rest	0.0	Lung	0.0
Macrophages LPS	0.0	Thymus	0.9
HUVEC none	0.0	Kidney	7.0
HUVEC starved	0.7		

General_screening_panel_v1.5 Summary: Ag4970 Expression of this gene is almost exclusive to small intestine (CT=31.2). Thus, expression of this gene could be used to differentiate between this sample and other samples on this panel.

Panel 4.1D Summary: Ag4970 Significant expression of this gene is detected in a liver cirrhosis sample (CT = 30.2). Furthermore, expression of this gene is not detected in normal liver in Panel 1.3D, suggesting that its expression is unique to liver cirrhosis. Therefore, therapeutic modulation of the expression or function of this gene may be used to diagnose this condition and to reduce or inhibit fibrosis that occurs in liver cirrhosis.

10 N. CG138751-01: CAMP INDUCIBLE 2 PROTEIN-LIKE-PROTEIN.

Expression of gene CG138751-01 was assessed using the primer-probe set Ag4971, described in Table NA. Results of the RTQ-PCR runs are shown in Tables NB, NC and ND.

Table NA. Probe Name Ag4971

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ggaagcctatcagtatcgtaa-3'	22	179	297
Probe	TET-5'-cggagcagatcaaaccatcaatgat-3'-TAMRA	26	224	298
Reverse	5'-cacatggtgtcattgagactgt-3'	22	254	299

Table NB. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag4971, Run 296465693	Tissue Name	Rel. Exp.(%) Ag4971, Run 296465693
110967 COPD-F	0.8	112427 Match Control Psoriasis-F	0.1
110980 COPD-F	0.7	112418 Psoriasis-M	0.7
110968 COPD-M	1.0	112723 Match Control Psoriasis-M	0.5
110977 COPD-M	1.4	112419 Psoriasis-M	0.9
110989 Emphysema-F	0.6	112424 Match Control Psoriasis-M	0.2
110992 Emphysema-F	0.4	112420 Psoriasis-M	1.6
110993 Emphysema-F	0.9	112425 Match Control Psoriasis-M	0.0
110994 Emphysema-F	0.5	104689 (MF) OA Bone- Backus	63.3
110995 Emphysema-F	1.6	104690 (MF) Adj "Normal" Bone-Backus	10.4
110996 Emphysema-F	0.3	104691 (MF) OA Synovium-Backus	39.0
110997 Asthma-M	0.2	104692 (BA) OA Cartilage-Backus	0.0
111001 Asthma-F	0.7	104694 (BA) OA Bone- Backus	100.0
111002 Asthma-F	1.0	104695 (BA) Adj "Normal" Bone-Backus	32.5
111003 Atopic Asthma- F	1.2	104696 (BA) OA Synovium-Backus	38.4
111004 Atopic Asthma- F	1.3	104700 (SS) OA Bone- Backus	8.9
111005 Atopic Asthma- F	0.7	104701 (SS) Adj "Normal" Bone-Backus	20.4
111006 Atopic Asthma- F	0.2	104702 (SS) OA Synovium-Backus	17.9
111417 Allergy-M	0.4	117093 OA Cartilage Rep7	1.0
112347 Allergy-M	0.0	112672 OA Bone5	0.1
112349 Normal Lung-F	0.0	112673 OA Synovium5	0.0
112357 Normal Lung-F	0.4	112674 OA Synovial Fluid cells5	0.0
112354 Normal Lung- M	0.0	117100 OA Cartilage Rep14	0.6
112374 Crohns-F	1.0	112756 OA Bone9	0.4

I12389 Match Control Crohns-F	2.6	I12757 OA Synovium9	0.2
I12375 Crohns-F	1.0	I12758 OA Synovial Fluid Cells9	0.4
I12732 Match Control Crohns-F	3.2	I17125 RA Cartilage Rep2	2.2
I12725 Crohns-M	0.1	I13492 Bone2 RA	2.2
I12387 Match Control Crohns-M	0.6	I13493 Synovium2 RA	0.3
I12378 Crohns-M	0.0	I13494 Syn Fluid Cells RA	1.5
I12390 Match Control Crohns-M	0.0	I13499 Cartilage4 RA	1.0
I12726 Crohns-M	0.9	I13500 Bone4 RA	1.2
I12731 Match Control Crohns-M	0.0	I13501 Synovium4 RA	0.6
I12380 Ulcer Col-F	0.3	I13502 Syn Fluid Cells4 RA	0.6
I12734 Match Control Ulcer Col-F	6.1	I13495 Cartilage3 RA	1.3
I12384 Ulcer Col-F	1.1	I13496 Bone3 RA	1.5
I12737 Match Control Ulcer Col-F	0.2	I13497 Synovium3 RA	0.9
I12386 Ulcer Col-F	0.7	I13498 Syn Fluid Cells3 RA	2.0
I12738 Match Control Ulcer Col-F	1.1	I17106 Normal Cartilage Rep20	0.5
I12381 Ulcer Col-M	0.0	I13663 Bone3 Normal	0.0
I12735 Match Control Ulcer Col-M	0.1	I13664 Synovium3 Normal	0.0
I12382 Ulcer Col-M	2.3	I13665 Syn Fluid Cells3 Normal	0.0
I12394 Match Control Ulcer Col-M	0.3	I17107 Normal Cartilage Rep22	0.2
I12383 Ulcer Col-M	1.8	I13667 Bone4 Normal	0.1
I12736 Match Control Ulcer Col-M	2.7	I13668 Synovium4 Normal	0.2
I12423 Psoriasis-F	0.4	I13669 Syn Fluid Cells4 Normal	0.2

Table NC. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag4971, Run 228926585	Tissue Name	Rel. Exp.(%) Ag4971, Run 228926585
Adipose	4.0	Renal ca. TK-10	1.0
Melanoma* Hs688(A).T	0.0	Bladder	4.9

Melanoma* Hs688(B).T	0.4	Gastric ca. (liver met.) NCI-N87	1.4
Melanoma* M14	19.2	Gastric ca. KATO III	0.2
Melanoma* LOXIMVI	0.1	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	67.4	Colon ca. SW480	1.8
Squamous cell carcinoma SCC-4	24.8	Colon ca.* (SW480 met) SW620	0.1
Testis Pool	0.7	Colon ca. HT29	0.1
Prostate ca.* (bone met) PC-3	6.2	Colon ca. HCT-116	2.5
Prostate Pool	0.9	Colon ca. CaCo-2	0.5
Placenta	4.2	Colon cancer tissue	9.7
Uterus Pool	0.6	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.5
Ovarian ca. SK-OV-3	1.2	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.7
Ovarian ca. OVCAR-5	3.3	Small Intestine Pool	0.3
Ovarian ca. IGROV-1	0.0	Stomach Pool	0.5
Ovarian ca. OVCAR-8	0.1	Bone Marrow Pool	0.5
Ovary	0.8	Fetal Heart	0.2
Breast ca. MCF-7	0.3	Heart Pool	0.4
Breast ca. MDA-MB-231	43.5	Lymph Node Pool	0.7
Breast ca. BT 549	0.1	Fetal Skeletal Muscle	0.7
Breast ca. T47D	0.1	Skeletal Muscle Pool	0.2
Breast ca. MDA-N	0.7	Spleen Pool	13.8
Breast Pool	0.4	Thymus Pool	0.6
Trachea	9.7	CNS cancer (glio/astro) U87-MG	2.2
Lung	0.0	CNS cancer (glio/astro) U-118-MG	25.5
Fetal Lung	0.6	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF- 539	0.5
Lung ca. LX-1	0.1	CNS cancer (astro) SNB- 75	12.7
Lung ca. NCI-H146	0.3	CNS cancer (glio) SNB- 19	0.0
Lung ca. SHP-77	0.1	CNS cancer (glio) SF-295	11.3
Lung ca. A549	0.0	Brain (Amygdala) Pool	0.8
Lung ca. NCI-H526	0.0	Brain (cerebellum)	1.1
Lung ca. NCI-H23	0.0	Brain (fetal)	0.7
Lung ca. NCI-H460	0.1	Brain (Hippocampus) Pool	0.7
Lung ca. HOP-62	2.0	Cerebral Cortex Pool	0.2

Lung ca. NCI-H522	0.3	Brain (Substantia nigra) Pool	0.8
Liver	2.2	Brain (Thalamus) Pool	0.4
Fetal Liver	3.4	Brain (whole)	1.5
Liver ca. HepG2	2.3	Spinal Cord Pool	0.4
Kidney Pool	1.7	Adrenal Gland	100.0
Fetal Kidney	0.1	Pituitary gland Pool	0.1
Renal ca. 786-0	0.3	Salivary Gland	59.0
Renal ca. A498	0.1	Thyroid (female)	0.8
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	57.8
Renal ca. UO-31	0.6	Pancreas Pool	1.0

Table ND. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4971, Run 223692675	Tissue Name	Rel. Exp.(%) Ag4971, Run 223692675
Secondary Th1 act	0.2	HUVEC IL-1beta	0.1
Secondary Th2 act	0.4	HUVEC IFN gamma	0.2
Secondary Tr1 act	0.9	HUVEC TNF alpha + IFN gamma	0.1
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.3
Secondary Th2 rest	0.7	HUVEC IL-11	0.1
Secondary Tr1 rest	0.2	Lung Microvascular EC none	0.0
Primary Th1 act	0.1	Lung Microvascular EC TNFalpha + IL-1beta	2.2
Primary Th2 act	0.4	Microvascular Dermal EC none	0.4
Primary Tr1 act	0.1	Microvascular Dermal EC TNFalpha + IL-1beta	0.9
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	4.8
Primary Th2 rest	0.0	Small airway epithelium none	5.4
Primary Tr1 rest	0.1	Small airway epithelium TNFalpha + IL-1beta	3.9
CD45RA CD4 lymphocyte act	0.1	Coronary artery SMC rest	0.2
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.2
CD8 lymphocyte act	0.2	Astrocytes rest	0.1
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	0.1
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.1
CD4 lymphocyte none	0.1	KU-812 (Basophil) PMA/ionomycin	3.0

2ry Th1/Th2/Tr1 _anti-CD95 CH11	0.3	CCD1106 (Keratinocytes) none	39.5
LAK cells rest	45.1	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	28.5
LAK cells IL-2	0.1	Liver cirrhosis	0.4
LAK cells IL-2+IL-12	0.4	NCI-H292 none	15.4
LAK cells IL-2+IFN gamma	0.4	NCI-H292 IL-4	10.3
LAK cells IL-2+ IL-18	0.2	NCI-H292 IL-9	21.8
LAK cells PMA/ionomycin	17.8	NCI-H292 IL-13	8.1
NK Cells IL-2 rest	0.1	NCI-H292 IFN gamma	17.7
Two Way MLR 3 day	4.6	HPAEC none	0.1
Two Way MLR 5 day	1.7	HPAEC TNF alpha + IL-1 beta	0.4
Two Way MLR 7 day	0.4	Lung fibroblast none	0.2
PBMC rest	3.1	Lung fibroblast TNF alpha + IL-1 beta	0.2
PBMC PWM	0.1	Lung fibroblast IL-4	0.1
PBMC PHA-L	0.3	Lung fibroblast IL-9	0.3
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.5
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.1
B lymphocytes PWM	0.2	Dermal fibroblast CCD1070 rest	0.2
B lymphocytes CD40L and IL-4	0.6	Dermal fibroblast CCD1070 TNF alpha	0.9
EOL-1 dbcAMP	2.6	Dermal fibroblast CCD1070 IL-1 beta	0.3
EOL-1 dbcAMP PMA/ionomycin	1.9	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	59.9	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	21.9	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	100.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	32.3	Neutrophils rest	0.7
Monocytes LPS	5.9	Colon	1.5
Macrophages rest	30.6	Lung	7.3
Macrophages LPS	8.7	Thymus	1.0
HUVEC none	0.3	Kidney	0.1
HUVEC starved	0.7		

AI_comprehensive panel_v1.0 Summary: Ag4971 Highest expression of this gene is detected in orthoarthritis (OA) bone (CT=26.7). High to moderate levels of expression of this gene is also seen in OA and adjacent normal bone and OA synovium. In addition, moderate to low levels of expression of this gene is also seen in bone, cartilage,

synovium and synovial fluid samples derived from rheumatoid arthritis patient, OA cartilage, as well as, in samples derived from COPD lung, emphysema, atopic asthma, asthma, allergy, Crohn's disease (normal matched control and diseased), ulcerative colitis (normal matched control and diseased), and psoriasis (normal matched control and diseased).

- 5 Therefore, therapeutic modulation of this gene product may ameliorate symptoms/conditions associated with autoimmune and inflammatory disorders including psoriasis, allergy, asthma, inflammatory bowel disease, rheumatoid arthritis and osteoarthritis.

General_screening_panel_v1.5 Summary: Ag4971 Highest expression of this gene is detected in adrenal gland (CT=27.8). Moderate to low levels of expression of this gene is also seen in tissues with metabolic/endocrine function such as pancreas, adipose, thyroid, and liver. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

15 Moderate levels of expression of this gene is also seen in number of cancer cell lines derived from melanoma, pancreatic, brain, colon, breast and prostate cancers. Therefore, expression of this gene may be used as diagnostic marker to detect the presence of these cancers. Furthermore, therapeutic modulation of this gene may be useful in the treatment of these cancers.

20 In addition, low levels of expression of this gene is also seen in whole and fetal brain, amygdala, cerebellum and substantia nigra. Therefore, therapeutic modulation of this gene product may be useful in the treatment of central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

25 **Panel 4.1D Summary:** Ag4971 Highest expression of this gene is detected in anti-CD40 treated dendritic cells (CT=29). Moderate levels of expression of this gene is detected in dendritic cells, monocytes, macrophages, LAK cells, keratinocytes and mucoepidermoid NCI-H292 cells. Moderate to low levels of expression of this gene is also seen in PMA/ionomycin activated LAK cells, two way MLR, PBMC, eosinophils, small airway epithelium, TNFalpha + IL-1beta activated bronchial epithelium and microvascular dermal epithelium and lung. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead

to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

A. CG139363-01 and CG139363-02: Transmembrane protein

5 HTMP10-like protein.

Expression of gene CG139363-01 and CG139363-02 was assessed using the primer-probe set Ag4978, described in Table OA. Results of the RTQ-PCR runs are shown in Tables OB, OC and OD. Note that CG139363-02 represents a full-length physical clone.

Table OA. Probe Name Ag4978

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ggccctctcttgattagc-3'	19	134	300
Probe	TET-5'- cacagccctgctggtggctttactat -3'-TAMRA	26	177	301
Reverse	5'-cttcttcttcggtgaatcaaag- 3'	22	206	302

10 Table OB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4978, Run 224757409	Tissue Name	Rel. Exp.(%) Ag4978, Run 224757409
AD 1 Hippo	4.1	Control (Path) 3 Temporal Ctx	4.4
AD 2 Hippo	8.7	Control (Path) 4 Temporal Ctx	20.4
AD 3 Hippo	2.6	AD 1 Occipital Ctx	6.2
AD 4 Hippo	9.1	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	11.3	AD 3 Occipital Ctx	1.8
AD 6 Hippo	29.9	AD 4 Occipital Ctx	24.7
Control 2 Hippo	43.8	AD 5 Occipital Ctx	40.3
Control 4 Hippo	11.3	AD 6 Occipital Ctx	23.7
Control (Path) 3 Hippo	4.9	Control 1 Occipital Ctx	4.4
AD 1 Temporal Ctx	7.2	Control 2 Occipital Ctx	43.8
AD 2 Temporal Ctx	23.2	Control 3 Occipital Ctx	20.0
AD 3 Temporal Ctx	2.5	Control 4 Occipital Ctx	9.0
AD 4 Temporal Ctx	33.0	Control (Path) 1 Occipital Ctx	100.0
AD 5 Inf Temporal Ctx	40.3	Control (Path) 2 Occipital Ctx	15.8

AD 5 Sup Temporal Ctx	17.3	Control (Path) 3 Occipital Ctx	4.4
AD 6 Inf Temporal Ctx	44.4	Control (Path) 4 Occipital Ctx	13.3
AD 6 Sup Temporal Ctx	34.9	Control 1 Parietal Ctx	7.7
Control 1 Temporal Ctx	5.1	Control 2 Parietal Ctx	23.7
Control 2 Temporal Ctx	41.8	Control 3 Parietal Ctx	19.3
Control 3 Temporal Ctx	23.3	Control (Path) 1 Parietal Ctx	57.8
Control 3 Temporal Ctx	12.8	Control (Path) 2 Parietal Ctx	20.4
Control (Path) 1 Temporal Ctx	49.7	Control (Path) 3 Parietal Ctx	4.1
Control (Path) 2 Temporal Ctx	31.4	Control (Path) 4 Parietal Ctx	25.9

Table OC. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag4978, Run 228940920	Tissue Name	Rel. Exp.(%) Ag4978, Run 228940920
Adipose	0.0	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	0.0
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca. * (SW480 met) SW620	0.0
Testis Pool	0.0	Colon ca. HT29	0.0
Prostate ca. * (bone met) PC-3	0.0	Colon ca. HCT-116	0.0
Prostate Pool	0.0	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.0
Ovarian ca. OVCAR-5	0.0	Small Intestine Pool	0.0
Ovarian ca. IGROV-1	0.0	Stomach Pool	0.0
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.0
Ovary	0.0	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart.Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.0

Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.0
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	0.0
Breast Pool	0.0	Thymus Pool	2.8
Trachea	0.1	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	0.0	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF- 539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB- 75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB- 19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	62.9
Lung ca. NCI-H526	0.0	Brain (cerebellum)	25.9
Lung ca. NCI-H23	0.0	Brain (fetal)	0.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	51.8
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	66.0
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	48.6
Liver	0.0	Brain (Thalamus) Pool	100.0
Fetal Liver	0.0	Brain (whole)	64.6
Liver ca. HepG2	0.0	Spinal Cord Pool	24.3
Kidney Pool	0.0	Adrenal Gland	0.0
Fetal Kidney	0.0	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.0	Pancreas Pool	0.0

Table OD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4978, Run 223693384	Tissue Name	Rel. Exp.(%) Ag4978, Run 223693384
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0

Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.6
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCDI106 (Keratinocytes) none	0.0
LAK cells rest	1.0	CCDI106 (Keratinocytes) TNFalpha + IL-1beta	0.7
LAK cells IL-2	0.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.5
Two Way MLR 3 day	0.0	HPAEC none	0.6
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0

B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	0.0	Colon	0.0
Macrophages rest	0.0	Lung	0.5
Macrophages LPS	0.0	Thymus	100.0
HUVEC none	0.0	Kidney	0.6
HUVEC starved	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4978 This panel does not show differential expression of this gene in Alzheimer's disease. However, this profile confirms the expression of this gene at moderate levels in the brain. See Panel 1.5 for discussion of this gene in the central nervous system.

- 5 **General_screening_panel_v1.5 Summary:** Ag4978 Highest expression of this gene is seen in the thalamus (CT=26.7). Overall, expression of this gene appears to be highly associated with the brain. High levels of expression are seen in all regions of the CNS examined, including the hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or
- 10 function of this gene may be useful in the treatment of neurological disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

- Panel 4.1D Summary:** Ag4978 This transcript is expressed at significant levels only in the thymus (CT = 30.2). The putative protein encoded by this gene could therefore
- 15 play an important role in T cell development. Therapeutic modulation of the expression or function of this gene may modulate immune function (T cell development) and be important for organ transplant, AIDS treatment or post chemotherapy immune reconstitution.

P. CG140188-01: DC2-Like Protein.

Expression of gene CGI40188-01 was assessed using the primer-probe set Ag7417, described in Table PA. Results of the RTQ-PCR runs are shown in Table PB.

Table PA. Probe Name Ag7417

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - cattggctctatgactgatgaac- 3'	23	194	303
Probe	TET-5' - ccaagaaagctactggcctctgat- 3' -TAMRA	24	223	304
Reverse	5' - ggatgcaagtccttcataata-3'	22	269	305

Table PB. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag7417, Run 305065593	Tissue Name	Rel. Exp.(%) Ag7417, Run 305065593
Secondary Th1 act	21.0	HUVEC IL-1beta	48.0
Secondary Th2 act	29.5	HUVEC IFN gamma	37.9
Secondary Tr1 act	14.6	HUVEC TNF alpha + IFN gamma	18.8
Secondary Th1 rest	1.2	HUVEC TNF alpha + IL4	24.8
Secondary Th2 rest	3.6	HUVEC IL-11	22.2
Secondary Tr1 rest	4.2	Lung Microvascular EC none	100.0
Primary Th1 act	3.4	Lung Microvascular EC TNFalpha + IL-1beta	41.8
Primary Th2 act	24.0	Microvascular Dermal EC none	9.5
Primary Tr1 act	20.6	Microvascular Dermal EC TNFalpha + IL-1beta	18.7
Primary Th1 rest	1.4	Bronchial epithelium TNFalpha + IL1beta	22.4
Primary Th2 rest	2.5	Small airway epithelium none	9.2
Primary Tr1 rest	1.2	Small airway epithelium TNFalpha + IL-1beta	14.9
CD45RA CD4 lymphocyte act	22.8	Coronary artery SMC rest	49.0
CD45RO CD4 lymphocyte act	21.3	Coronary artery SMC TNFalpha + IL-1beta	45.1
CD8 lymphocyte act	13.8	Astrocytes rest	10.7
Secondary CD8 lymphocyte rest	2.2	Astrocytes TNFalpha + IL- 1beta	24.3
Secondary CD8 lymphocyte act	6.3	KU-812 (Basophil) rest	22.8

CD4 lymphocyte none	1.6	KU-812 (Basophil) PMA/ionomycin	31.4
2ry Th1/Th2/Tr1_anti- CD95 CH11	3.2	CCD1106 (Keratinocytes) none	23.8
LAK cells rest	6.5	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	10.7
LAK cells IL-2	4.2	Liver cirrhosis	11.3
LAK cells IL-2+IL-12	1.4	NCI-H292 none	25.2
LAK cells IL-2+IFN gamma	5.8	NCI-H292 IL-4	17.3
LAK cells IL-2+ IL-18	2.6	NCI-H292 IL-9	33.2
LAK cells PMA/ionomycin	9.5	NCI-H292 IL-13	20.6
NK Cells IL-2 rest	20.9	NCI-H292 IFN gamma	6.8
Two Way MLR 3 day	4.8	HPAEC none	15.2
Two Way MLR 5 day	2.5	HPAEC TNF alpha + IL-1 beta	54.3
Two Way MLR 7 day	4.3	Lung fibroblast none	22.2
PBMC rest	1.5	Lung fibroblast TNF alpha + IL-1 beta	21.0
PBMC PWM	5.1	Lung fibroblast IL-4	27.5
PBMC PHA-L	3.6	Lung fibroblast IL-9	30.1
Ramos (B cell) none	36.3	Lung fibroblast IL-13	15.2
Ramos (B cell) ionomycin	59.0	Lung fibroblast IFN gamma	38.4
B lymphocytes PWM	3.9	Dermal fibroblast CCD1070 rest	40.1
B lymphocytes CD40L and IL-4	10.4	Dermal fibroblast CCD1070 TNF alpha	49.7
EOL-1 dbcAMP	26.6	Dermal fibroblast CCD1070 IL-1 beta	31.9
EOL-1 dbcAMP PMA/ionomycin	14.1	Dermal fibroblast IFN gamma	17.0
Dendritic cells none	10.2	Dermal fibroblast IL-4	28.7
Dendritic cells LPS	5.1	Dermal Fibroblasts rest	8.4
Dendritic cells anti-CD40	3.5	Neutrophils TNFa+LPS	0.0
Monocytes rest	5.4	Neutrophils rest	1.2
Monocytes LPS	22.7	Colon	0.7
Macrophages rest	5.6	Lung	1.4
Macrophages LPS	3.7	Thymus	4.1
HUVEC none	43.8	Kidney	6.0
HUVEC starved	38.2		

CNS_neurodegeneration_v1.0 Summary: Ag7417 Expression of this gene is low/undetectable in all samples on this panel (CTs>35).

Panel 4.1D Summary: Ag7417 Highest expression of this gene is seen in untreated lung microvascular endothelial cells (CT=30.8). This gene is also expressed at moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, endothelial cell, basophil, astrocyte, monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues as well as in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

**Q. CG140305-01: COMPLEMENT-clq TUMOR NECROSIS
FACTOR-RELATED PROTEIN-LIKE PROTEIN.**

Expression of gene CG140305-01 was assessed using the primer-probe set Ag6486, described in Table QA. Results of the RTQ-PCR runs are shown in Tables QB, QC and QD.

Table QA. Probe Name Ag6486

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - tgctggatgtatctgatttgc - 3'	21	581	306
Probe	TET-5' - caacacagtcttccagcatgtacagct -3' - TAMRA	26	543	307
Reverse	5' - gtatgtgtaccttatgcacaatgg - 3'	24	519	308

Table QB. General_screening_panel_v1.6

Tissue Name	Rel. Exp.(%) Ag6486, Run 277240051	Tissue Name	Rel. Exp.(%) Ag6486, Run 277240051
Adipose	13.9	Renal ca. TK-10	12.4
Melanoma* Hs688(A).T	35.4	Bladder	14.7
Melanoma* Hs688(B).T	55.9	Gastric ca. (liver met.) NCI-N87	10.8
Melanoma* M14	1.4	Gastric ca. KATO III	0.3
Melanoma* LOXIMV1	1.3	Colon ca. SW-948	0.4
Melanoma* SK-MEL-5	2.0	Colon ca. SW480	0.6

Squamous cell carcinoma SCC-4	1.7	Colon ca.* (SW480 met) SW620	4.4
Testis Pool	30.8	Colon ca. HT29	2.6
Prostate ca.* (bone met) PC-3	2.8	Colon ca. HCT-116	4.9
Prostate Pool	15.2	Colon ca. CaCo-2	9.0
Placenta	2.6	Colon cancer tissue	33.7
Uterus Pool	3.8	Colon ca. SW1116	1.7
Ovarian ca. OVCAR-3	4.4	Colon ca. Colo-205	2.0
Ovarian ca. SK-OV-3	14.0	Colon ca. SW-48	1.0
Ovarian ca. OVCAR-4	1.0	Colon Pool	9.1
Ovarian ca. OVCAR-5	9.1	Small Intestine Pool	22.8
Ovarian ca. IGROV-1	4.6	Stomach Pool	15.5
Ovarian ca. OVCAR-8	1.2	Bone Marrow Pool	4.6
Ovary	3.1	Fetal Heart	10.5
Breast ca. MCF-7	7.2	Heart Pool	3.9
Breast ca. MDA-MB-231	5.0	Lymph Node Pool	5.8
Breast ca. BT 549	4.2	Fetal Skeletal Muscle	39.5
Breast ca. T47D	0.5	Skeletal Muscle Pool	1.8
Breast ca. MDA-N	1.1	Spleen Pool	4.4
Breast Pool	11.0	Thymus Pool	14.3
Trachea	26.8	CNS cancer (glio/astro) U87-MG	2.2
Lung	4.7	CNS cancer (glio/astro) U-118-MG	4.5
Fetal Lung	34.9	CNS cancer (neuro;met) SK-N-AS	3.8
Lung ca. NCI-N417	0.3	CNS cancer (astro) SF-539	2.2
Lung ca. LX-1	6.8	CNS cancer (astro) SNB-75	8.8
Lung ca. NCI-H146	12.9	CNS cancer (glio) SNB-19	3.7
Lung ca. SHP-77	19.8	CNS cancer (glio) SF-295	7.6
Lung ca. A549	0.7	Brain (Amygdala) Pool	9.1
Lung ca. NCI-H526	2.5	Brain (cerebellum)	100.0
Lung ca. NCI-H23	9.4	Brain (fetal)	20.2
Lung ca. NCI-H460	1.0	Brain (Hippocampus) Pool	13.5
Lung ca. HOP-62	4.9	Cerebral Cortex Pool	10.8
Lung ca. NCI-H522	1.2	Brain (Substantia nigra) Pool	7.4
Liver	0.4	Brain (Thalamus) Pool	14.7
Fetal Liver	5.7	Brain (whole)	8.0
Liver ca. HepG2	5.2	Spinal Cord Pool	24.8

Kidney Pool	20.2	Adrenal Gland	3.2
Fetal Kidney	65.1	Pituitary gland Pool	2.2
Renal ca. 786-0	10.4	Salivary Gland	18.7
Renal ca. A498	2.1	Thyroid (female)	2.1
Renal ca. ACHN	2.2	Pancreatic ca. CAPAN2	5.6
Renal ca. UO-31	1.2	Pancreas Pool	3.3

Table QC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag6486, Run 269282929	Tissue Name	Rel. Exp.(%) Ag6486, Run 269282929
Secondary Th1 act	15.8	HUVEC IL-1beta	9.7
Secondary Th2 act	27.9	HUVEC IFN gamma	8.8
Secondary Tr1 act	17.0	HUVEC TNF alpha + IFN gamma	7.5
Secondary Th1 rest	9.5	HUVEC TNF alpha + IL4	5.3
Secondary Th2 rest	9.4	HUVEC IL-11	8.5
Secondary Tr1 rest	8.8	Lung Microvascular EC none	92.0
Primary Th1 act	3.8	Lung Microvascular EC TNFalpha + IL-1beta	5.8
Primary Th2 act	38.2	Microvascular Dermal EC none	9.4
Primary Tr1 act	31.0	Microvascular Dermal EC TNFalpha + IL-1beta	8.5
Primary Th1 rest	9.5	Bronchial epithelium TNFalpha + IL1beta	5.4
Primary Th2 rest	13.2	Small airway epithelium none	0.0
Primary Tr1 rest	1.4	Small airway epithelium TNFalpha + IL-1beta	6.4
CD45RA CD4 lymphocyte act	34.6	Coronary artery SMC rest	3.9
CD45RO CD4 lymphocyte act	40.3	Coronary artery SMC TNFalpha + IL-1beta	5.4
CD8 lymphocyte act	21.9	Astrocytes rest	8.0
Secondary CD8 lymphocyte rest	6.9	Astrocytes TNFalpha + IL- 1beta	0.0
Secondary CD8 lymphocyte act	5.2	KU-812 (Basophil) rest	52.1
CD4 lymphocyte none	13.5	KU-812 (Basophil) PMA/ionomycin	33.2
2ry Th1/Th2/Tr1_anti- CD95 CH11	24.1	CCD1106 (Keratinocytes) none	8.0
LAK cells rest	8.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	5.7
LAK cells IL-2	10.4	Liver cirrhosis	20.7
LAK cells IL-2+IL-12	1.9	NCI-H292 none	34.6

LAK cells IL-2+IFN gamma	14.3	NCI-H292 IL-4	24.3
LAK cells IL-2+ IL-18	21.0	NCI-H292 IL-9	34.4
LAK cells PMA/ionomycin	7.0	NCI-H292 IL-13	29.9
NK Cells IL-2 rest	70.2	NCI-H292 IFN gamma	17.2
Two Way MLR 3 day	23.2	HPAEC none	7.0
Two Way MLR 5 day	2.0	HPAEC TNF alpha + IL-1 beta	8.0
Two Way MLR 7 day	6.9	Lung fibroblast none	5.0
PBMC rest	1.6	Lung fibroblast TNF alpha + IL-1 beta	9.6
PBMC PWM	7.2	Lung fibroblast IL-4	0.0
PBMC PHA-L	15.2	Lung fibroblast IL-9	5.3
Ramos (B cell) none	9.7	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	17.2	Lung fibroblast IFN gamma	14.1
B lymphocytes PWM	6.8	Dermal fibroblast CCD1070 rest	12.9
B lymphocytes CD40L and IL-4	56.6	Dermal fibroblast CCD1070 TNF alpha	100.0
EOL-1 dbcAMP	50.0	Dermal fibroblast CCD1070 IL-1 beta	8.0
EOL-1 dbcAMP PMA/ionomycin	8.1	Dermal fibroblast IFN gamma	9.2
Dendritic cells none	13.2	Dermal fibroblast IL-4	28.1
Dendritic cells LPS	3.1	Dermal Fibroblasts rest	7.9
Dendritic cells anti-CD40	3.6	Neutrophils TNFa+LPS	2.9
Monocytes rest	0.0	Neutrophils rest	6.2
Monocytes LPS	7.7	Colon	46.0
Macrophages rest	2.4	Lung	9.0
Macrophages LPS	0.0	Thymus	51.4
HUVEC none	1.7	Kidney	78.5
HUVEC starved	22.7		

Table QD. Panel CNS_1.1

Tissue Name	Rel. Exp.(%) Ag6486, Run 271481506	Tissue Name	Rel. Exp.(%) Ag6486, Run 271481506
Cing Gyr Depression2	26.8	BA17 PSP2	5.6
Cing Gyr Depression	13.8	BA17 PSP	11.3
Cing Gyr PSP2	4.8	BA17 Huntington's2	17.2
Cing Gyr PSP	52.9	BA17 Huntington's	20.0
Cing Gyr Huntington's2	33.2	BA17 Parkinson's2	26.4
Cing Gyr Huntington's	53.2	BA17 Parkinson's	31.4

Cing Gyr Parkinson's2	0.0	BA17 Alzheimer's2	7.5
Cing Gyr Parkinson's	53.6	BA17 Control2	12.6
Cing Gyr Alzheimer's2	14.6	BA17 Control	19.8
Cing Gyr Alzheimer's	23.5	BA9 Depression2	4.9
Cing Gyr Control2	16.6	BA9 Depression	0.4
Cing Gyr Control	52.9	BA9 PSP2	6.0
Temp Pole Depression2	15.6	BA9 PSP	15.6
Temp Pole PSP2	0.0	BA9 Huntington's2	28.3
Temp Pole PSP	2.1	BA9 Huntington's	36.3
Temp Pole Huntington's	28.3	BA9 Parkinson's2	26.4
Temp Pole Parkinson's2	31.4	BA9 Parkinson's	32.8
Temp Pole Parkinson's	16.2	BA9 Alzheimer's2	4.6
Temp Pole Alzheimer's2	5.9	BA9 Alzheimer's	1.8
Temp Pole Alzheimer's	4.7	BA9 Control2	59.9
Temp Pole Control2	33.9	BA9 Control	14.3
Temp Pole Control	4.0	BA7 Depression	17.4
Glob Palladus Depression	12.1	BA7 PSP2	12.1
Glob Palladus PSP2	7.0	BA7 PSP	14.5
Glob Palladus PSP	14.7	BA7 Huntington's2	76.8
Glob Palladus Parkinson's2	30.4	BA7 Huntington's	26.1
Glob Palladus Parkinson's	73.2	BA7 Parkinson's2	17.4
Glob Palladus Alzheimer's2	7.4	BA7 Parkinson's	19.3
Glob Palladus Alzheimer's	9.1	BA7 Alzheimer's2	2.8
Glob Palladus Control2	6.7	BA7 Control2	14.4
Glob Palladus Control	19.5	BA7 Control	9.9
Sub Nigra Depression2	9.3	BA4 Depression2	6.7
Sub Nigra Depression	4.5	BA4 Depression	13.5
Sub Nigra PSP2	17.4	BA4 PSP2	10.2
Sub Nigra Huntington's2	46.0	BA4 PSP	7.5
Sub Nigra Huntington's	63.7	BA4 Huntington's2	13.5
Sub Nigra Parkinson's2	76.3	BA4 Huntington's	14.5
Sub Nigra Alzheimer's2	27.4	BA4 Parkinson's2	34.2
Sub Nigra Control2	36.9	BA4 Parkinson's	39.2
Sub Nigra Control	100.0	BA4 Alzheimer's2	2.3

BA17 Depression2	25.3	BA4 Control2	22.5
BA17 Depression	19.2	BA4 Control	15.0

General_screening_panel_v1.6 Summary: Ag6486 Highest expression of this gene is detected in brain cerebellum (CT=27.8). In addition, moderate levels of expression of this gene is also seen in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, therapeutic modulation of this gene product may be useful in the treatment of central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Moderate levels of expression of this gene is also seen in cluster of cancer cell lines derived from pancreatic, gastric, colon, lung, liver, renal, breast, ovarian, prostate, squamous cell carcinoma, melanoma and brain cancers. Thus, therapeutic modulation of the expression or function of this gene may be effective in the treatment of pancreatic, gastric, colon, lung, liver, renal, breast, ovarian, prostate, squamous cell carcinoma, melanoma and brain cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, fetal liver and the gastrointestinal tract. This gene encodes a splice variant of the complement C1q tumor necrosis factor-related protein, a member of the C1q family. This family includes proteins such as complement subunit C1q, adiponectin, gliacolin, C1q-related protein, cerebellin, CORS26 etc., all of which are secreted. These proteins have been implicated in tissue differentiation, immune regulation, energy homeostasis, synaptic function and in diseases such as obesity, diabetes and neurodegeneration. Adiponectin, a member of C1q family and protein closely related to complement C1q tumor necrosis factor-related protein, is induced over 100-fold in adipocyte differentiation (Scherer *et al.*, 1995, J Biol Chem 270(45):26746-9 PMID: 7592907) and is involved in adipocyte signaling (Hu *et al.*, 1996, J Biol Chem 271(18):10697-703 PMID: 8631877). Recently, adiponectin has been shown to reverse insulin resistance in mouse models of lipoatrophy and obesity (Yamauchi *et al.*, 2001, Nat Med 7(8):941-6 PMID: 11479627). Therefore this protein, and proteins related to it, are potential antigens for development protein therapeutics for use in the treatment of obesity and type II diabetes.

This gene is expressed at much higher levels in fetal (CTs=29-32) when compared to adult skeletal muscle, lung and liver (CTs=32-35.9). This observation suggests that expression of this gene can be used to distinguish fetal from adult skeletal muscle, lung and liver. In addition, the relative overexpression of this gene in fetal tissue suggests that the protein product may enhance growth or development of these tissues in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the protein encoded by this gene could be useful in treatment of muscle, lung and liver related diseases.

Panel 4.1D Summary: Ag6486 Highest expression of this gene is detected in TNF alpha treated dermal fibroblast (CT=32.4). In addition, moderate to low levels of expression of this gene is also seen in activated T cells, IL-2 treated NK Cells, CD40L and IL-4 treated B lymphocytes, eosinophils, lung microvascular endothelial cells, basophils, NCI-H292 mucoepidermoid cells, and normal tissues represented by colon, thymus and kidney. Therefore, therapeutic modulation of the activity of this gene or its protein product, through the use of protein therapeutics or antibodies, might be beneficial in the treatment of autoimmune and inflammatory diseases that involve these cell and tissue types, such as lupus erythematosus, asthma, emphysema, Crohn's disease, ulcerative colitis, rheumatoid arthritis, osteoarthritis, and psoriasis.

Panel CNS_1.1 Summary: Ag6486 This panel confirms the expression of this gene at low levels in the brains of an independent group of individuals. See Panel 1.6 for a discussion of this gene in treatment of central nervous system disorders.

R. CG140639-01 and CG140639-02: Flotillin-2 (Reggie-1) (REG-1)-like protein.

Expression of gene CG140639-01 and CG140639-02 was assessed using the primer-probe set Ag5036, described in Table RA. Results of the RTQ-PCR runs are shown in Tables RB and RC. Note that CG140639-02 represents a full-length physical clone.

Table RA. Probe Name Ag5036

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gggtaagaatgtgcaggacat-3'	21	349	309
Probe	TET-5'-aaaacgtcgtcctgcagaccctg-3'-TAMRA	23	372	310

Reverse	5' - tgataaatctgtccactgtca-3'	22	426	311
---------	----------------------------------	----	-----	-----

Table RB. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag5036, Run 228967203	Tissue Name	Rel. Exp.(%) Ag5036, Run 228967203
Adipose	10.4	Renal ca. TK-10	44.4
Melanoma* Hs688(A).T	23.0	Bladder	33.4
Melanoma* Hs688(B).T	18.4	Gastric ca. (liver met.) NCI-N87	25.9
Melanoma* M14	45.4	Gastric ca. KATO III	41.5
Melanoma* LOXIMVI	13.9	Colon ca. SW-948	14.5
Melanoma* SK-MEL-5	23.2	Colon ca. SW480	61.6
Squamous cell carcinoma SCC-4	6.1	Colon ca.* (SW480 met) SW620	47.3
Testis Pool	6.3	Colon ca. HT29	37.1
Prostate ca.* (bone met) PC-3	15.4	Colon ca. HCT-116	39.5
Prostate Pool	17.2	Colon ca. CaCo-2	68.3
Placenta	33.7	Colon cancer tissue	20.3
Uterus Pool	11.0	Colon ca. SW1116	8.9
Ovarian ca. OVCAR-3	57.0	Colon ca. Colo-205	12.1
Ovarian ca. SK-OV-3	100.0	Colon ca. SW-48	11.7
Ovarian ca. OVCAR-4	27.5	Colon Pool	15.3
Ovarian ca. OVCAR-5	39.8	Small Intestine Pool	11.3
Ovarian ca. IGROV-1	44.4	Stomach Pool	8.2
Ovarian ca. OVCAR-8	26.1	Bone Marrow Pool	4.8
Ovary	9.3	Fetal Heart	16.4
Breast ca. MCF-7	24.8	Heart Pool	9.9
Breast ca. MDA-MB-231	89.5	Lymph Node Pool	12.7
Breast ca. BT 549	47.6	Fetal Skeletal Muscle	11.0
Breast ca. T47D	16.2	Skeletal Muscle Pool	22.2
Breast ca. MDA-N	11.7	Spleen Pool	14.6
Breast Pool	10.7	Thymus Pool	10.9
Trachea	22.7	CNS cancer (glio/astro) U87-MG	39.8
Lung	2.2	CNS cancer (glio/astro) U-118-MG	23.0
Fetal Lung	29.9	CNS cancer (neuro;met) SK-N-AS	19.6
Lung ca. NCI-N417	6.6	CNS cancer (astro) SF- 539	13.9

Lung ca. LX-1	68.8	CNS cancer (astro) SNB-75	34.4
Lung ca. NCI-H146	13.4	CNS cancer (glio) SNB-19	43.8
Lung ca. SHP-77	41.2	CNS cancer (glio) SF-295	48.0
Lung ca. A549	52.1	Brain (Amygdala) Pool	18.2
Lung ca. NCI-H526	23.7	Brain (cerebellum)	47.0
Lung ca. NCI-H23	44.1	Brain (fetal)	27.5
Lung ca. NCI-H460	29.1	Brain (Hippocampus) Pool	15.8
Lung ca. HOP-62	43.8	Cerebral Cortex Pool	23.5
Lung ca. NCI-H522	36.9	Brain (Substantia nigra) Pool	18.6
Liver	5.8	Brain (Thalamus) Pool	22.8
Fetal Liver	47.6	Brain (whole)	23.2
Liver ca. HepG2	15.3	Spinal Cord Pool	11.7
Kidney Pool	20.9	Adrenal Gland	13.8
Fetal Kidney	8.7	Pituitary gland Pool	2.7
Renal ca. 786-0	21.9	Salivary Gland	14.1
Renal ca. A498	19.1	Thyroid (female)	11.3
Renal ca. ACHN	50.0	Pancreatic ca. CAPAN2	21.8
Renal ca. UO-31	39.2	Pancreas Pool	18.2

Table RC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag5036, Run 223740995	Tissue Name	Rel. Exp.(%) Ag5036, Run 223740995
Secondary Th1 act	42.9	HUVEC IL-1beta	27.2
Secondary Th2 act	54.3	HUVEC IFN gamma	42.0
Secondary Tr1 act	35.4	HUVEC TNF alpha + IFN gamma	21.8
Secondary Th1 rest	21.2	HUVEC TNF alpha + IL4	31.9
Secondary Th2 rest	35.6	HUVEC IL-11	32.1
Secondary Tr1 rest	20.7	Lung Microvascular EC none	72.2
Primary Th1 act	13.3	Lung Microvascular EC TNFalpha + IL-1beta	36.6
Primary Th2 act	34.6	Microvascular Dermal EC none	54.7
Primary Tr1 act	37.1	Microvascular Dermal EC TNFalpha + IL-1beta	26.6
Primary Th1 rest	19.5	Bronchial epithelium TNFalpha + IL1beta	25.5
Primary Th2 rest	23.0	Small airway epithelium none	14.2
Primary Tr1 rest	40.1	Small airway epithelium TNFalpha + IL-1beta	26.2

CD45RA CD4 lymphocyte act	34.6	Coronary artery SMC rest	16.2
CD45RO CD4 lymphocyte act	51.8	Coronary artery SMC TNFalpha + IL-1 beta	19.3
CD8 lymphocyte act	28.3	Astrocytes rest	14.4
Secondary CD8 lymphocyte rest	31.0	Astrocytes TNFalpha + IL-1 beta	12.9
Secondary CD8 lymphocyte act	17.3	KU-812 (Basophil) rest	50.0
CD4 lymphocyte none	16.4	KU-812 (Basophil) PMA/ionomycin	67.4
2ry Th1/Th2/Tr1_anti-CD95 CH11	39.0	CCD1106 (Keratinocytes) none	31.4
LAK cells rest	32.5	CCD1106 (Keratinocytes) TNFalpha + IL-1 beta	53.2
LAK cells IL-2	38.7	Liver cirrhosis	10.4
LAK cells IL-2+IL-12	11.1	NCI-H292 none	44.4
LAK cells IL-2+IFN gamma	13.2	NCI-H292 IL-4	69.7
LAK cells IL-2+ IL-18	21.0	NCI-H292 IL-9	66.9
LAK cells PMA/ionomycin	11.3	NCI-H292 IL-13	57.0
NK Cells IL-2 rest	49.3	NCI-H292 IFN gamma	49.3
Two Way MLR 3 day	27.9	HPAEC none	29.7
Two Way MLR 5 day	22.1	HPAEC TNF alpha + IL-1 beta	25.2
Two Way MLR 7 day	28.7	Lung fibroblast none	42.3
PBMC rest	20.6	Lung fibroblast TNF alpha + IL-1 beta	28.3
PBMC PWM	26.1	Lung fibroblast IL-4	25.3
PBMC PHA-L	34.9	Lung fibroblast IL-9	31.0
Ramos (B cell) none	19.8	Lung fibroblast IL-13	22.1
Ramos (B cell) ionomycin	35.1	Lung fibroblast IFN gamma	40.1
B lymphocytes PWM	21.9	Dermal fibroblast CCD1070 rest	19.6
B lymphocytes CD40L and IL-4	51.8	Dermal fibroblast CCD1070 TNF alpha	65.1
EOL-1 dbcAMP	19.1	Dermal fibroblast CCD1070 IL-1 beta	12.1
EOL-1 dbcAMP PMA/ionomycin	11.2	Dermal fibroblast IFN gamma	17.1
Dendritic cells none	27.5	Dermal fibroblast IL-4	22.1
Dendritic cells LPS	25.5	Dermal Fibroblasts rest	26.8
Dendritic cells anti-CD40	26.4	Neutrophils TNFa+LPS	16.5
Monocytes rest	53.2	Neutrophils rest	100.0
Monocytes LPS	31.4	Colon	6.5

Macrophages rest	27.2	Lung	22.4
Macrophages LPS	16.5	Thymus	13.5
HUVEC none	17.1	Kidney	25.2
HUVEC starved	38.4		

General_screening_panel_v1.5 Summary: Ag5036 Highest expression of this gene is seen in an ovarian cancer cell line (CT=27). This gene is widely expressed in this panel, with moderate expression seen in brain, colon, gastric, lung, breast, ovarian, and melanoma cancer cell lines. This gene encodes a protein with homology to flotillin-2, an
5 integral membrane protein of the plasmalemmal microdomains involved in vesicular trafficking and signal transduction. Cho has suggested that this molecule is involved in cell adhesion (Genomics 27: 251-258, 1995.). Thus, based on this expression profile and the homology of this gene to flotillin, this protein product may be involved in cell survival and/or proliferation. Modulation of this gene product may be useful in the treatment of
10 cancer.

Among tissues with metabolic function, this gene is expressed at moderate levels in pituitary, adipose, adrenal gland, pancreas, thyroid, and adult and fetal skeletal muscle, heart, and liver. Flotillin-2 may play a role in the glucose uptake pathway (Baumann, Nature 2000 Sep 14;407(6801):202-7). This widespread expression among these metabolic tissues
15 and the homology to flotillin suggest that this gene product may play a role in normal neuroendocrine and metabolic function and that dysregulated expression of this gene may contribute to neuroendocrine disorders or metabolic diseases, such as obesity and diabetes.

In addition, this gene is expressed at much higher levels in fetal liver tissue (CT=28) when compared to expression in the adult counterpart (CT=31). Thus, expression of this
20 gene may be used to differentiate between the fetal and adult source of this tissue.

This gene is also expressed at moderate levels in the CNS, including the hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurologic disorders, such as Alzheimer's disease, Parkinson's disease,
25 schizophrenia, multiple sclerosis, stroke and epilepsy.

Panel 4.1D Summary: Ag5036 Highest expression of this gene is seen in neutrophils (CT=28.2). This gene is also expressed at moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include

members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

S. CG140843-01: INTEGRIN BETA-5 PRECURSOR PROTEIN-LIKE PROTEIN.

Expression of gene CG140843-01 was assessed using the primer-probe set Ag7404, described in Table SA. Results of the RTQ-PCR runs are shown in Table SB. Table SA.
Probe Name Ag7404

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctgcatggggagggtcaa-3'	17	852	312
Probe	TET-5'- aagtaccaacacccaactgacgtctc -3'-TAMRA	26	873	313
Reverse	5'-gctggggcactcaaagact-3'	19	907	314

Table SB. General_screening_panel_v1.6

Tissue Name	Rel. Exp.(%) Ag7404, Run 306066735	Tissue Name	Rel. Exp.(%) Ag7404, Run 306066735
Adipose	6.4	Renal ca. TK-10	38.2
Melanoma* Hs688(A).T	21.6	Bladder	0.0
Melanoma* Hs688(B).T	14.2	Gastric ca. (liver met.) NCI-N87	34.2
Melanoma* M14	9.5	Gastric ca. KATO III	16.6
Melanoma* LOXIMVI	2.9	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	9.3	Colon ca. SW480	100.0
Squamous cell carcinoma SCC-4	2.9	Colon ca.* (SW480 met) SW620	11.0
Testis Pool	5.1	Colon ca. HT29	17.7

Prostate ca.* (bone met) PC-3	8.2	Colon ca. HCT-116	18.2
Prostate Pool	3.4	Colon ca. CaCo-2	16.4
Placenta	0.0	Colon cancer tissue	7.6
Uterus Pool	10.2	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	50.0	Colon ca. Colo-205	4.3
Ovarian ca. SK-OV-3	27.4	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	26.1
Ovarian ca. OVCAR-5	40.1	Small Intestine Pool	11.3
Ovarian ca. IGROV-1	4.2	Stomach Pool	18.7
Ovarian ca. OVCAR-8	6.5	Bone Marrow Pool	4.5
Ovary	13.3	Fetal Heart	3.3
Breast ca. MCF-7	24.1	Heart Pool	10.3
Breast ca. MDA-MB-231	46.3	Lymph Node Pool	20.4
Breast ca. BT 549	10.6	Fetal Skeletal Muscle	0.0
Breast ca. T47D	7.4	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	5.6	Spleen Pool	15.5
Breast Pool	23.2	Thymus Pool	6.7
Trachea	9.0	CNS cancer (glio/astro) U87-MG	10.2
Lung	9.4	CNS cancer (glio/astro) U-118-MG	18.8
Fetal Lung	11.4	CNS cancer (neuro;met) SK-N-AS	30.8
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF- 539	17.8
Lung ca. LX-1	21.3	CNS cancer (astro) SNB- 75	43.8
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB- 19	6.8
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	20.7
Lung ca. A549	35.4	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	5.6
Lung ca. NCI-H23	2.8	Brain (fetal)	0.0
Lung ca. NCI-H460	10.7	Brain (Hippocampus) Pool	7.3
Lung ca. HOP-62	12.8	Cerebral Cortex Pool	3.2
Lung ca. NCI-H522	7.1	Brain (Substantia nigra) Pool	7.1
Liver	0.0	Brain (Thalamus) Pool	3.3
Fetal Liver	0.0	Brain (whole)	3.4
Liver ca. HepG2	19.9	Spinal Cord Pool	13.4
Kidney Pool	9.0	Adrenal Gland	6.6
Fetal Kidney	12.2	Pituitary gland Pool	0.0
Renal ca. 786-0	16.5	Salivary Gland	0.0

Renal ca. A498	3.6	Thyroid (female)	0.0
Renal ca. ACHN	13.4	Pancreatic ca. CAPAN2	35.6
Renal ca. UO-31	19.6	Pancreas Pool	4.2

CNS_neurodegeneration_v1.0 Summary: Ag7404 Expression of this gene is low/undetectable in all samples on this panel (CTs>35).

General_screening_panel_v1.6 Summary: Ag7404 Expression of this gene is restricted to a sample derived from a colon cancer cell line (CT=34.8). Thus, expression of this gene could be used to differentiate between this sample and other samples on this panel and as a marker to detect the presence of colon cancer. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of colon cancer.

T. CG141540-01: IL1 receptor -type-2-like protein

Expression of gene CG141540-01 was assessed using the primer-probe sets Ag5237 and Ag5236, described in Tables TA and TB. Results of the RTQ-PCR runs are shown in Tables TC, TD and TE.

Table TA. Probe Name Ag5237

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agatggtctgactgtgctatg-3'	21	1143	315
Probe	TET-5'- tcacatcaagactttcaatcctatccc a-3'-TAMRA	29	1167	316
Reverse	5'-gaattatttcattccattttatttc- 3'	24	1199	317

Table TB. Probe Name Ag5236

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-acgcatcaagaggtaagact- 3'	21	744	318
Probe	TET-5'- ccggcacacccttaaccacat- 3'-TAMRA	22	794	319
Reverse	5'-gtgtcattggcgtcca-3'	17	823	320

Table TC. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag5236, Run 229545061	Tissue Name	Rel. Exp.(%) Ag5236, Run 229545061
110967 COPD-F	0.0	112427 Match Control Psoriasis-F	0.0
110980 COPD-F	0.0	112418 Psoriasis-M	0.0

I10968 COPD-M	1.4	I12723 Match Control Psoriasis-M	0.0
I10977 COPD-M	0.0	I12419 Psoriasis-M	0.0
I10989 Emphysema-F	0.0	I12424 Match Control Psoriasis-M	1.6
I10992 Emphysema-F	2.7	I12420 Psoriasis-M	3.8
I10993 Emphysema-F	1.8	I12425 Match Control Psoriasis-M	0.0
I10994 Emphysema-F	0.0	I04689 (MF) OA Bone-Backus	3.4
I10995 Emphysema-F	11.4	I04690 (MF) Adj "Normal" Bone-Backus	1.7
I10996 Emphysema-F	6.1	I04691 (MF) OA Synovium-Backus	1.7
I10997 Asthma-M	4.6	I04692 (BA) OA Cartilage-Backus	0.0
I11001 Asthma-F	0.0	I04694 (BA) OA Bone-Backus	1.8
I11002 Asthma-F	0.0	I04695 (BA) Adj "Normal" Bone-Backus	0.9
I11003 Atopic Asthma-F	0.7	I04696 (BA) OA Synovium-Backus	0.0
I11004 Atopic Asthma-F	0.0	I04700 (SS) OA Bone-Backus	4.4
I11005 Atopic Asthma-F	0.0	I04701 (SS) Adj "Normal" Bone-Backus	3.0
I11006 Atopic Asthma-F	0.0	I04702 (SS) OA Synovium-Backus	1.5
I11417 Allergy-M	0.0	I17093 OA Cartilage Rep7	1.4
I12347 Allergy-M	0.0	I12672 OA Bone5	0.0
I12349 Normal Lung-F	0.0	I12673 OA Synovium5	0.0
I12357 Normal Lung-F	0.0	I12674 OA Synovial Fluid cells5	1.2
I12354 Normal Lung-M	0.0	I17100 OA Cartilage Rep14	0.0
I12374 Crohns-F	0.0	I12756 OA Bone9	0.0
I12389 Match Control Crohns-F	6.1	I12757 OA Synovium9	0.7
I12375 Crohns-F	0.0	I12758 OA Synovial Fluid Cells9	1.5
I12732 Match Control Crohns-F	30.8	I17125 RA Cartilage Rep2	0.0
I12725 Crohns-M	0.0	I13492 Bone2 RA	0.9
I12387 Match Control Crohns-M	1.5	I13493 Synovium2 RA	0.0

112378 Crohns-M	0.0	113494 Syn Fluid Cells RA	2.0
112390 Match Control Crohns-M	0.0	113499 Cartilage4 RA	2.5
112726 Crohns-M	1.1	113500 Bone4 RA	2.2
112731 Match Control Crohns-M	1.7	113501 Synovium4 RA	1.9
112380 Ulcer Col-F	0.0	113502 Syn Fluid Cells4 RA	0.0
112734 Match Control Ulcer Col-F	100.0	113495 Cartilage3 RA	4.5
112384 Ulcer Col-F	1.8	113496 Bone3 RA	5.3
112737 Match Control Ulcer Col-F	0.6	113497 Synovium3 RA	2.3
112386 Ulcer Col-F	1.2	113498 Syn Fluid Cells3 RA	2.0
112738 Match Control Ulcer Col-F	4.6	117106 Normal Cartilage Rep20	0.0
112381 Ulcer Col-M	0.0	113663 Bone3 Normal	0.0
112735 Match Control Ulcer Col-M	0.0	113664 Synovium3 Normal	0.0
112382 Ulcer Col-M	7.9	113665 Syn Fluid Cells3 Normal	0.0
112394 Match Control Ulcer Col-M	0.0	117107 Normal Cartilage Rep22	0.0
112383 Ulcer Col-M	1.9	113667 Bone4 Normal	0.0
112736 Match Control Ulcer Col-M	0.0	113668 Synovium4 Normal	0.0
112423 Psoriasis-F	2.3	113669 Syn Fluid Cells4 Normal	0.0

Table TD. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag5236, Run 237228536	Tissue Name	Rel. Exp.(%) Ag5236, Run 237228536
Adipose	22.2	Renal ca. TK-10	3.2
Melanoma* Hs688(A).T	0.0	Bladder	16.8
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	11.4
Melanoma* M14	0.0	Gastric ca. KATO III	40.1
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	29.3
Melanoma* SK-MEL-5	0.7	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	8.2	Colon ca.* (SW480 met) SW620	4.9
Testis Pool	1.4	Colon ca. HT29	13.3
Prostate ca.* (bone met) PC-3	1.0	Colon ca. HCT-116	0.0

Prostate Pool	0.0	Colon ca. CaCo-2	42.0
Placenta	13.4	Colon cancer tissue	35.1
Uterus Pool	3.5	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	26.2	Colon ca. Colo-205	67.4
Ovarian ca. SK-OV-3	100.0	Colon ca. SW-48	3.3
Ovarian ca. OVCAR-4	80.7	Colon Pool	2.7
Ovarian ca. OVCAR-5	1.3	Small Intestine Pool	2.9
Ovarian ca. IGROV-1	25.3	Stomach Pool	2.0
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	5.0
Ovary	18.6	Fetal Heart	0.8
Breast ca. MCF-7	0.0	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.0
Breast ca. BT 549	3.2	Fetal Skeletal Muscle	0.5
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	59.5
Breast Pool	1.2	Thymus Pool	17.4
Trachea	12.6	CNS cancer (glio/astro) U87-MG	6.8
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	6.6	CNS cancer (neuro;met) SK-N-AS	6.3
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF- 539	1.1
Lung ca. LX-1	2.1	CNS cancer (astro) SNB- 75	18.7
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB- 19	85.3
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	3.4
Lung ca. A549	8.0	Brain (Amygdala) Pool	2.1
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	0.0	Brain (fetal)	3.4
Lung ca. NCI-H460	3.4	Brain (Hippocampus) Pool	0.0
Lung ca. HOP-62	2.7	Cerebral Cortex Pool	0.0
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.0
Liver	0.8	Brain (Thalamus) Pool	0.8
Fetal Liver	4.1	Brain (whole)	1.8
Liver ca. HepG2	7.0	Spinal Cord Pool	0.0
Kidney Pool	2.8	Adrenal Gland	3.2
Fetal Kidney	0.0	Pituitary gland Pool	0.0
Renal ca. 786-0	11.3	Salivary Gland	1.0
Renal ca. A498	3.3	Thyroid (female)	0.8
Renal ca. ACHN	0.6	Pancreatic ca. CAPAN2	0.0

Renal ca. UO-31	0.0	Pancreas Pool	13.9
-----------------	-----	---------------	------

Table TE. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag5236, Run 229788311	Tissue Name	Rel. Exp.(%) Ag5236, Run 229788311
Secondary Th1 act	0.0	HUVEC IL-1 beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1 beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1 beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1 beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1 beta	5.9
CD45RA CD4 lymphocyte act	2.1	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	1.1	Coronary artery SMC TNFalpha + IL-1 beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.7
Secondary CD8 lymphocyte rest	1.4	Astrocytes TNFalpha + IL- 1 beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.9
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti- CD95 CH11	0.0	CCD1106 (Keratinocytes) none	2.5
LAK cells rest	3.7	CCD1106 (Keratinocytes) TNFalpha + IL-1 beta	7.4
LAK cells IL-2	0.0	Liver cirrhosis	2.5
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0

LAK cells PMA/ionomycin	12.4	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	2.2	HPAEC none	0.0
Two Way MLR 5 day	1.1	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.9	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.9	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	4.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	1.6	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	10.3	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	4.7	Neutrophils TNFa+LPS	85.3
Monocytes rest	0.0	Neutrophils rest	100.0
Monocytes LPS	1.0	Colon	0.0
Macrophages rest	0.0	Lung	1.0
Macrophages LPS	0.0	Thymus	2.8
HUVEC none	0.0	Kidney	0.0
HUVEC starved	0.0		

AI_comprehensive_panel_v1.0 Summary: Ag5236 Expression of this gene is limited to a normal tissue sample adjacent to Crohn's and normal tissue sample adjacent to ulcerative colitis (CTs=32-34). Thus, expression of this gene could be used to differentiate between these samples and other samples on this panel.

- 5 **General_screening_panel_v1.5 Summary:** Ag5236 Highest expression of this gene is seen in an ovarian cancer cell line (CT=32). Low but significant levels of expression are also seen in clusters of cell lines derived from brain, ovarian, colon and gastric cancers. Thus, this gene product may be involved in these cancers. Low levels of expression are also seen in adipose and pancreas suggesting a role for this gene product in the pathogenesis of
- 10 metabolic disorders including obesity and diabetes.

Panel 4.1D Summary: Ag5236 This gene is expressed exclusively in neutrophils. Thus, expression of this gene could be used to differentiate between these samples and other samples on this panel and as a marker of neutrophils.

U. CG141580-01: KIAA 1467 protein-like protein.

- 5 Expression of gene CG141580-01 was assessed using the primer-probe set Ag7248, described in Table UA. Results of the RTQ-PCR runs are shown in Tables UB and UC.

Table UA. Probe Name Ag7248

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - gtctatgactaggaacattttgtgtac -3'	29	2255	321
Probe	TET-5' - ccacaacactaaaatatacacacacacag c-3' -TAMRA	30	2289	322
Reverse	5' - cttaggacatacctggaaaataacttc- 3'	27	2320	323

Table UB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag7248, Run 296423801	Tissue Name	Rel. Exp.(%) Ag7248, Run 296423801
AD 1 Hippo	6.3	Control (Path) 3 Temporal Ctx	1.2
AD 2 Hippo	14.1	Control (Path) 4 Temporal Ctx	12.9
AD 3 Hippo	2.5	AD 1 Occipital Ctx	6.1
AD 4 Hippo	1.9	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	100.0	AD 3 Occipital Ctx	1.2
AD 6 Hippo	27.2	AD 4 Occipital Ctx	9.9
Control 2 Hippo	13.3	AD 5 Occipital Ctx	25.5
Control 4 Hippo	1.8	AD 6 Occipital Ctx	23.5
Control (Path) 3 Hippo	2.5	Control 1 Occipital Ctx	0.9
AD 1 Temporal Ctx	3.5	Control 2 Occipital Ctx	43.8
AD 2 Temporal Ctx	8.2	Control 3 Occipital Ctx	7.6
AD 3 Temporal Ctx	1.7	Control 4 Occipital Ctx	1.2
AD 4 Temporal Ctx	17.8	Control (Path) 1 Occipital Ctx	66.4
AD 5 Inf Temporal Ctx	60.7	Control (Path) 2 Occipital Ctx	7.6

AD 5 Sup Temporal Ctx	20.7	Control (Path) 3 Occipital Ctx	0.5
AD 6 Inf Temporal Ctx	25.5	Control (Path) 4 Occipital Ctx	6.7
AD 6 Sup Temporal Ctx	30.8	Control 1 Parietal Ctx	1.8
Control 1 Temporal Ctx	1.3	Control 2 Parietal Ctx	16.3
Control 2 Temporal Ctx	24.5	Control 3 Parietal Ctx	14.4
Control 3 Temporal Ctx	5.5	Control (Path) 1 Parietal Ctx	67.8
Control 3 Temporal Ctx	2.8	Control (Path) 2 Parietal Ctx	14.6
Control (Path) 1 Temporal Ctx	37.9	Control (Path) 3 Parietal Ctx	1.7
Control (Path) 2 Temporal Ctx	21.0	Control (Path) 4 Parietal Ctx	35.6

Table UC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag7248, Run 296417628	Tissue Name	Rel. Exp.(%) Ag7248, Run 296417628
Secondary Th1 act	53.6	HUVEC IL-1beta	36.1
Secondary Th2 act	50.0	HUVEC IFN gamma	37.6
Secondary Tr1 act	16.8	HUVEC TNF alpha + IFN gamma	7.1
Secondary Th1 rest	1.7	HUVEC TNF alpha + IL4	14.3
Secondary Th2 rest	2.0	HUVEC IL-11	11.3
Secondary Tr1 rest	6.7	Lung Microvascular EC none	100.0
Primary Th1 act	5.6	Lung Microvascular EC TNFalpha + IL-1 beta	12.3
Primary Th2 act	33.7	Microvascular Dermal EC none	18.2
Primary Tr1 act	27.7	Microvascular Dermal EC TNFalpha + IL-1 beta	8.0
Primary Th1 rest	1.2	Bronchial epithelium TNFalpha + IL1 beta	11.0
Primary Th2 rest	2.1	Small airway epithelium none	11.7
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1 beta	27.9
CD45RA CD4 lymphocyte act	36.9	Coronary artery SMC rest	34.9
CD45RO CD4 lymphocyte act	55.5	Coronary artery SMC TNFalpha + IL-1 beta	41.5
CD8 lymphocyte act	11.1	Astrocytes rest	10.9
Secondary CD8 lymphocyte rest	5.2	Astrocytes TNFalpha + IL-1 beta	10.0

Secondary CD8 lymphocyte act	4.5	KU-812 (Basophil) rest	63.3
CD4 lymphocyte none	4.3	KU-812 (Basophil) PMA/ionomycin	81.2
2ry Th1/Th2/Tr1_anti-CD95 CH11	4.7	CCD1106 (Keratinocytes) none	28.1
LAK cells rest	20.4	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	3.9
LAK cells IL-2	4.6	Liver cirrhosis	5.3
LAK cells IL-2+IL-12	1.6	NCI-H292 none	13.5
LAK cells IL-2+IFN gamma	9.1	NCI-H292 IL-4	20.6
LAK cells IL-2+ IL-18	3.9	NCI-H292 IL-9	21.6
LAK cells PMA/ionomycin	37.1	NCI-H292 IL-13	28.7
NK Cells IL-2 rest	12.8	NCI-H292 IFN gamma	14.5
Two Way MLR 3 day	27.9	HPAEC none	10.8
Two Way MLR 5 day	3.3	HPAEC TNF alpha + IL-1 beta	35.8
Two Way MLR 7 day	8.8	Lung fibroblast none	44.4
PBMC rest	5.4	Lung fibroblast TNF alpha + IL-1 beta	45.4
PBMC PWM	11.4	Lung fibroblast IL-4	12.5
PBMC PHA-L	12.4	Lung fibroblast IL-9	14.6
Ramos (B cell) none	20.6	Lung fibroblast IL-13	11.3
Ramos (B cell) ionomycin	53.2	Lung fibroblast IFN gamma	37.6
B lymphocytes PWM	4.9	Dermal fibroblast CCD1070 rest	24.8
B lymphocytes CD40L and IL-4	17.9	Dermal fibroblast CCD1070 TNF alpha	33.0
EOL-1 dbcAMP	38.4	Dermal fibroblast CCD1070 IL-1 beta	21.0
EOL-1 dbcAMP PMA/ionomycin	37.1	Dermal fibroblast IFN gamma	16.2
Dendritic cells none	24.0	Dermal fibroblast IL-4	56.6
Dendritic cells LPS	3.4	Dermal Fibroblasts rest	22.7
Dendritic cells anti-CD40	2.6	Neutrophils TNFa+LPS	1.5
Monocytes rest	8.8	Neutrophils rest	2.9
Monocytes LPS	41.5	Colon	5.4
Macrophages rest	11.6	Lung	1.1
Macrophages LPS	3.0	Thymus	5.6
HUVEC none	22.1	Kidney	50.0
HUVEC starved	35.8		

CNS_neurodegeneration_v1.0 Summary: Ag7248 This panel confirms the expression of this gene at low levels in the brains of an independent group of individuals.

However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Low levels of expression of this gene in brain regions suggests that this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy,
 5 multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag7248 Highest expression of this gene is detected in lung microvascular endothelial cells (CT=32). Expression of this gene is down-regulated on activation of these endothelial cells by cytokines. Thus, this gene may play a role in the maintenance of the integrity of the microvasculature. Therefore, therapeutics designed for
 10 this putative protein could be beneficial for the treatment of diseases associated with damaged microvasculature including inflammatory diseases of lung, such as asthma, allergy, and chronic obstructive pulmonary diseases.

In addition, low to moderate levels of expression of this gene is also seen in lung and dermal fibroblasts, keratinocytes, basophils, coronary artery SMC, cytokine activated small
 15 airway epithelium, dermal microvascular EC, HUVEC, cytokine activated HPAEC, activated monocytes, eosinophils, Ramos B cells, two way MLR, activated LAK cells, and various types of activated T cells. Therefore, therapeutic modulation of this gene may be useful in the treatment of inflammatory and autoimmune diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and
 20 osteoarthritis.

V. CG141643-01:RIKEN 2010001CC9 protein-like protein.

Expression of gene CG141643-01 was assessed using the primer-probe set Ag5057, described in Table VA. Results of the RTQ-PCR runs are shown in Tables VB, VC and VD.

Table VA. Probe Name Ag5057

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gcgtccagggaaccttcttc-3'	19	355	324
Probe	TET-5'- actgggtcctgctggcactagctct- 3'-TAMRA	25	386	325
Reverse	5'-caacggacaagagcaggtt-3'	19	415	326

25

Table VB. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag5057, Run 219965745	Tissue Name	Rel. Exp.(%) Ag5057, Run 219965745
I10967 COPD-F	7.0	I12427 Match Control Psoriasis-F	80.7
I10980 COPD-F	3.6	I12418 Psoriasis-M	7.0
I10968 COPD-M	12.7	I12723 Match Control Psoriasis-M	5.0
I10977 COPD-M	0.0	I12419 Psoriasis-M	8.8
I10989 Emphysema-F	30.1	I12424 Match Control Psoriasis-M	25.0
I10992 Emphysema-F	22.2	I12420 Psoriasis-M	70.7
I10993 Emphysema-F	11.1	I12425 Match Control Psoriasis-M	48.3
I10994 Emphysema-F	5.8	I04689 (MF) OA Bone- Backus	7.0
I10995 Emphysema-F	98.6	I04690 (MF) Adj "Normal" Bone-Backus	8.1
I10996 Emphysema-F	13.6	I04691 (MF) OA Synovium-Backus	10.7
I10997 Asthma-M	19.9	I04692 (BA) OA Cartilage-Backus	10.4
I11001 Asthma-F	8.2	I04694 (BA) OA Bone- Backus	7.7
I11002 Asthma-F	21.3	I04695 (BA) Adj "Normal" Bone-Backus	8.8
I11003 Atopic Asthma- F	25.5	I04696 (BA) OA Synovium-Backus	3.0
I11004 Atopic Asthma- F	87.1	I04700 (SS) OA Bone- Backus	6.8
I11005 Atopic Asthma- F	32.1	I04701 (SS) Adj "Normal" Bone-Backus	10.7
I11006 Atopic Asthma- F	14.1	I04702 (SS) OA Synovium-Backus	9.6
I11417 Allergy-M	33.2	I17093 OA Cartilage Rep7	5.0
I12347 Allergy-M	10.7	I12672 OA Bone5	20.9
I12349 Normal Lung-F	25.3	I12673 OA Synovium5	6.4
I12357 Normal Lung-F	30.1	I12674 OA Synovial Fluid cells5	14.1
I12354 Normal Lung- M	17.2	I17100 OA Cartilage Rep14	8.0
I12374 Crohns-F	10.7	I12756 OA Bone9	23.0
I12389 Match Control Crohns-F	9.5	I12757 OA Synovium9	3.3
I12375 Crohns-F	14.2	I12758 OA Synovial Fluid Cells9	11.0

112732 Match Control Crohns-F	54.7	117125 RA Cartilage Rep2	2.9
112725 Crohns-M	11.8	113492 Bone2 RA	27.5
112387 Match Control Crohns-M	10.3	113493 Synovium2 RA	17.4
112378 Crohns-M	7.4	113494 Syn Fluid Cells RA	38.4
112390 Match Control Crohns-M	42.6	113499 Cartilage4 RA	49.3
112726 Crohns-M	32.3	113500 Bone4 RA	51.8
112731 Match Control Crohns-M	42.3	113501 Synovium4 RA	45.1
112380 Ulcer Col-F	12.6	113502 Syn Fluid Cells4 RA	34.4
112734 Match Control Ulcer Col-F	85.3	113495 Cartilage3 RA	13.8
112384 Ulcer Col-F	50.3	113496 Bone3 RA	23.8
112737 Match Control Ulcer Col-F	16.7	113497 Synovium3 RA	22.8
112386 Ulcer Col-F	2.6	113498 Syn Fluid Cells3 RA	24.5
112738 Match Control Ulcer Col-F	100.0	117106 Normal Cartilage Rep20	6.3
112381 Ulcer Col-M	3.5	113663 Bone3 Normal	5.8
112735 Match Control Ulcer Col-M	24.1	113664 Synovium3 Normal	7.5
112382 Ulcer Col-M	20.0	113665 Syn Fluid Cells3 Normal	7.1
112394 Match Control Ulcer Col-M	2.7	117107 Normal Cartilage Rep22	2.7
112383 Ulcer Col-M	23.2	113667 Bone4 Normal	8.2
112736 Match Control Ulcer Col-M	11.3	113668 Synovium4 Normal	14.7
112423 Psoriasis-F	11.5	113669 Syn Fluid Cells4 Normal	19.2

Table VC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag5057, Run 219514716	Tissue Name	Rel. Exp.(%) Ag5057, Run 219514716
Adipose	0.5	Renal ca. TK-10	1.6
Melanoma* Hs688(A).T	0.1	Bladder	12.9
Melanoma* Hs688(B).T	0.1	Gastric ca. (liver met.) NCI-N87	43.2
Melanoma* M14	0.4	Gastric ca. KATO III	100.0
Melanoma* LOXIMVI	0.2	Colon ca. SW-948	21.9
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	1.2

Squamous cell carcinoma SCC-4	15.8	Colon ca.* (SW480 met) SW620	0.4
Testis Pool	0.8	Colon ca. HT29	24.3
Prostate ca.* (bone met) PC-3	0.5	Colon ca. HCT-116	53.2
Prostate Pool	2.1	Colon ca. CaCo-2	26.2
Placenta	0.3	Colon cancer tissue	33.2
Uterus Pool	0.1	Colon ca. SW1116	14.1
Ovarian ca. OVCAR-3	1.7	Colon ca. Colo-205	29.5
Ovarian ca. SK-OV-3	3.3	Colon ca. SW-48	25.3
Ovarian ca. OVCAR-4	22.1	Colon Pool	0.5
Ovarian ca. OVCAR-5	24.1	Small Intestine Pool	1.5
Ovarian ca. IGROV-1	1.0	Stomach Pool	1.0
Ovarian ca. OVCAR-8	0.4	Bone Marrow Pool	0.1
Ovary	0.8	Fetal Heart	0.1
Breast ca. MCF-7	17.3	Heart Pool	0.2
Breast ca. MDA-MB-231	0.5	Lymph Node Pool	1.0
Breast ca. BT 549	0.5	Fetal Skeletal Muscle	0.1
Breast ca. T47D	51.1	Skeletal Muscle Pool	0.1
Breast ca. MDA-N	0.4	Spleen Pool	0.4
Breast Pool	1.2	Thymus Pool	1.8
Trachea	4.7	CNS cancer (glio/astro) U87-MG	0.8
Lung	0.9	CNS cancer (glio/astro) U-118-MG	1.0
Fetal Lung	1.6	CNS cancer (neuro;met) SK-N-AS	0.5
Lung ca. NCI-N417	0.2	CNS cancer (astro) SF- 539	0.5
Lung ca. LX-1	30.4	CNS cancer (astro) SNB- 75	0.6
Lung ca. NCI-H146	9.7	CNS cancer (glio) SNB- 19	0.9
Lung ca. SHP-77	0.3	CNS cancer (glio) SF-295	2.9
Lung ca. A549	0.9	Brain (Amygdala) Pool	0.2
Lung ca. NCI-H526	6.3	Brain (cerebellum)	0.9
Lung ca. NCI-H23	1.3	Brain (fetal)	0.8
Lung ca. NCI-H460	1.1	Brain (Hippocampus) Pool	0.2
Lung ca. HOP-62	0.7	Cerebral Cortex Pool	0.2
Lung ca. NCI-H522	0.9	Brain (Substantia nigra) Pool	0.2
Liver	0.7	Brain (Thalamus) Pool	0.3
Fetal Liver	1.4	Brain (whole)	0.2
Liver ca. HepG2	1.0	Spinal Cord Pool	0.3

Kidney Pool	0.9	Adrenal Gland	0.8
Fetal Kidney	1.0	Pituitary gland Pool	1.3
Renal ca. 786-0	0.5	Salivary Gland	1.4
Renal ca. A498	0.1	Thyroid (female)	2.1
Renal ca. ACHN	0.5	Pancreatic ca. CAPAN2	40.3
Renal ca. UO-31	0.5	Pancreas Pool	3.3

Table VD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag5057, Run 220366655	Tissue Name	Rel. Exp.(%) Ag5057, Run 220366655
Secondary Th1 act	14.0	HUVEC IL-1beta	3.8
Secondary Th2 act	15.8	HUVEC IFN gamma	9.7
Secondary Tr1 act	1.4	HUVEC TNF alpha + IFN gamma	6.7
Secondary Th1 rest	3.8	HUVEC TNF alpha + IL4	2.4
Secondary Th2 rest	5.3	HUVEC IL-11	2.3
Secondary Tr1 rest	10.2	Lung Microvascular EC none	17.1
Primary Th1 act	2.3	Lung Microvascular EC TNFalpha + IL-1beta	4.9
Primary Th2 act	8.2	Microvascular Dermal EC none	0.0
Primary Tr1 act	9.7	Microvascular Dermal EC TNFalpha + IL-1beta	4.4
Primary Th1 rest	14.5	Bronchial epithelium TNFalpha + IL1beta	24.3
Primary Th2 rest	0.0	Small airway epithelium none	19.9
Primary Tr1 rest	8.3	Small airway epithelium TNFalpha + IL-1beta	100.0
CD45RA CD4 lymphocyte act	0.6	Coronary artery SMC rest	0.9
CD45RO CD4 lymphocyte act	17.3	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	17.3	Astrocytes rest	4.2
Secondary CD8 lymphocyte rest	5.6	Astrocytes TNFalpha + IL- 1beta	7.1
Secondary CD8 lymphocyte act	15.1	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	5.0	KU-812 (Basophil) PMA/ionomycin	3.2
2ry Th1/Th2/Tr1_anti- CD95 CH11	7.3	CCD1106 (Keratinocytes) none	53.2
LAK cells rest	15.8	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	59.0
LAK cells IL-2	27.4	Liver cirrhosis	27.9
LAK cells IL-2+IL-12	7.7	NCI-H292 none	3.5

LAK cells IL-2+IFN gamma	9.6	NCI-H292 IL-4	11.8
LAK cells IL-2+ IL-18	6.3	NCI-H292 IL-9	3.7
LAK cells PMA/ionomycin	6.0	NCI-H292 IL-13	10.0
NK Cells IL-2 rest	20.2	NCI-H292 IFN gamma	20.3
Two Way MLR 3 day	12.2	HPAEC none	3.2
Two Way MLR 5 day	7.7	HPAEC TNF alpha + IL-1 beta	9.4
Two Way MLR 7 day	4.4	Lung fibroblast none	2.5
PBMC rest	2.4	Lung fibroblast TNF alpha + IL-1 beta	4.0
PBMC PWM	11.2	Lung fibroblast IL-4	4.3
PBMC PHA-L	9.5	Lung fibroblast IL-9	7.1
Ramos (B cell) none	5.8	Lung fibroblast IL-13	4.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	2.7
B lymphocytes PWM	4.2	Dermal fibroblast CCD1070 rest	4.5
B lymphocytes CD40L and IL-4	9.5	Dermal fibroblast CCD1070 TNF alpha	8.8
EOL-1 dbcAMP	5.0	Dermal fibroblast CCD1070 IL-1 beta	2.2
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.5
Dendritic cells none	17.8	Dermal fibroblast IL-4	2.4
Dendritic cells LPS	2.3	Dermal Fibroblasts rest	2.8
Dendritic cells anti-CD40	5.6	Neutrophils TNFa+LPS	2.4
Monocytes rest	7.6	Neutrophils rest	2.3
Monocytes LPS	13.2	Colon	39.2
Macrophages rest	23.7	Lung	16.5
Macrophages LPS	2.6	Thymus	19.1
HUVEC none	9.3	Kidney	44.8
HUVEC starved	13.8		

AI_comprehensive panel_v1.0 Summary: Ag5057 Highest expression of this gene is detected in a matched control for ulcerative colitis (CT=30.2). This gene shows a ubiquitous expression with moderate to low levels of expression in normal and diseased lung (COPD, emphysema and asthma), normal and diseased colon (Crohn's and ulcerative colitis), psoriasis, bone, cartilage, synovium and synovial fluids from normal and patients suffering from orthoarthritis and rheumatoid arthritis. Therefore, therapeutic modulation of this gene may be useful in the treatment of autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General_screening_panel_v1.4 Summary: Ag5057 This gene is expressed at a high to moderate level in pancreatic, gastric, colon cancer and some breast and ovarian cancer cell line with the highest expression seen in a gastric cancer cell line (KATO III, CT=26.33). It is also expressed at a low level in lung, CNS and prostate cancer cell lines as well as most of the normal tissues on this panel. Hence it may be used as a marker to differentiate cancer cells from normal tissue and therapeutic modulation of the gene product can be used for the treatment of these cancers.

In addition, low levels of expression of this gene is also seen in some regions of central nervous system including fetal brain, cerebellum, thalamus and spinal cord. Therefore, therapeutic modulation of this gene product may be useful in the treatment of central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Among tissues with metabolic or endocrine function, this gene is expressed at low levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

Panel 4.1D Summary: Ag5057 Highest expression of this gene is detected in TNF alpha and IL-1 beta treated small airway epithelium (CT=31.7). Expression of this gene is enhanced in cytokine treated small airway epithelium as compared to the resting cells (CT=34). Therefore, modulation of the expression or activity of the protein encoded by this transcript through the application of small molecule therapeutics may be useful in the treatment of asthma, COPD, and emphysema.

Moderate to low levels of expression of this gene is also seen in activated secondary polarized T cells, activated memory T cells, CD8 lymphocytes, resting and IL-2 treated LAK cells, IL-2 treated NK cells, dendritic cells, resting macrophage, activated monocytes, starved HUVEC cells, activated bronchial epithelium, keratinocytes, liver cirrhosis, activated NCI-H292 cells, and normal tissues represented by colon, lung, thymus and kidney. Therefore, therapeutic modulation of this gene may be useful in the treatment of autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

W. CG142003-01: Plasma Protease C1 Inhibitor Precursor Protein-like Protein.

Expression of gene CG142003-01 was assessed using the primer-probe set Ag5686, described in Table WA. Note that CG142003-01 represents a full-length physical clone.

5 Table WA. Probe Name Ag5686

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - catcgcagaaacctgaagatc - 3'	21	187	327
Probe	TET-5' - taccactgatgaaccaccacacaaac - 3' - TAMRA	26	225	328
Reverse	5' - cagccacaaaaataacagctaa - 3'	22	251	329

AI_comprehensive_panel_v1.0 Summary: Ag5686 Expression of this gene is low/undetectable in all samples on this panel (CTs>35).

General_screening_panel_v1.5 Summary: Ag5686 Expression of this gene is low/undetectable in all samples on this panel (CTs>35).

10 **Panel 4.1D Summary:** Ag5686 Expression of this gene is low/undetectable in all samples on this panel (CTs>35).

X. CG142023-01: 6230421J19Rik protein-like protein

Expression of gene CG142023-01 was assessed using the primer-probe set Ag7414, described in Table XA.

15 Table XA. Probe Name Ag7414

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - gaagagcatcgccaccat - 3'	18	798	330
Probe	TET-5' - ccctgggctctatcatttactgtgt - 3' - TAMRA	25	887	331
Reverse	5' - gctttctggtctccatgaactt - 3'	22	916	332

Y. CG142092-01: C4b-BINDING PROTEIN ALPHA CHAIN PRECURSOR PROTEIN-LIKE PROTEIN.

Expression of gene CG142092-01 was assessed using the primer-probe set Ag6869, described in Table YA. Results of the RTQ-PCR runs are shown in Tables YB and YC. Note that CG142092-01 represents a full-length physical clone.

Table YA. Probe Name Ag6869

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - tcacctacagctgtgaacaa - 3'	20	585	333
Probe	TET- 5' - caggcaaaagactcatgcagtgctctcc - 3' - TAMRA	27	612	334
Reverse	5' - ttttcacatcctctgggttt - 3'	20	640	335

5

Table YB. General_screening_panel_v1.6

Tissue Name	Rel. Exp.(%) Ag6869, Run 278387610	Tissue Name	Rel. Exp.(%) Ag6869, Run 278387610
Adipose	0.0	Renal ca. TK-10	9.8
Melanoma* Hs688(A).T	0.0	Bladder	33.4
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	0.0	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.0
Prostate Pool	0.0	Colon ca. CaCo-2	9.1
Placenta	0.0	Colon cancer tissue	2.0
Uterus Pool	0.3	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.9
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.0
Ovarian ca. OVCAR-5	2.8	Small Intestine Pool	0.3
Ovarian ca. IGROV-1	0.0	Stomach Pool	0.0
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.0
Ovary	0.6	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart Pool	0.2
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.2
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.5
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.5
Breast ca. MDA-N	0.0	Spleen Pool	0.0

Breast Pool	0.3	Thymus Pool	0.3
Trachea	0.6	CNS cancer (glio/astro) U87-MG	0.3
Lung	0.6	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	2.4	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF- 539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB- 75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB- 19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	1.2	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.7
Lung ca. NCI-H23	0.0	Brain (fetal)	0.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	0.0
Lung ca. HOP-62	1.3	Cerebral Cortex Pool	0.0
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.3
Liver	100.0	Brain (Thalamus) Pool	0.2
Fetal Liver	5.8	Brain (whole)	6.2
Liver ca. HepG2	16.2	Spinal Cord Pool	0.0
Kidney Pool	0.4	Adrenal Gland	0.0
Fetal Kidney	0.2	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.0	Pancreas Pool	1.1

Table YC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag6869, Run 310594482	Tissue Name	Rel. Exp.(%) Ag6869, Run 310594482
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0

Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	100.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0

EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	0.0	Colon	1.8
Macrophages rest	0.0	Lung	23.7
Macrophages LPS	0.0	Thymus	0.0
HUVEC none	0.0	Kidney	0.0
HUVEC starved	0.0		

General_screening_panel_v1.6 Summary: Ag6869 Highest expression of this gene is seen in liver (CT=30). In addition, this gene is expressed at much higher levels in adult liver when compared to expression in the fetal counterpart (CT=34). Thus, expression of this gene may be used to differentiate between the fetal and adult source of this tissue.

- 5 Low but significant levels of expression are also seen in cancer cell lines derived from liver, renal, and colon cancers, as well as in normal bladder and whole brain. This gene encodes a protein with homology to C4BP, a regulatory protein synthesized by the liver that is involved in the regulation of the classical pathway of complement and the natural anticoagulant pathway. Thus, the restricted pattern of expression of this protein, with
- 10 highest expression in the liver, is consistent with the its characterization as a novel C4BP.

- Panel 4.1D Summary:** Ag6869 Low expression of this gene is exclusively seen in liver cirrhosis sample (CT=34). The putative C4b-binding protein encoded for by this gene could potentially allow cells within the liver to respond to specific microenvironmental signals. Therefore, therapeutic modulation of this gene through the use of antibodies or
- 15 small molecule drug may potentially modulate liver function and play a role in the identification and treatment of inflammatory or autoimmune diseases which effect the liver including liver cirrhosis and fibrosis.

Z. CG142092-02: C4b-binding protein alpha-chain precursor protein-like protein.

- 20 Expression of gene CG142092-02 was assessed using the primer-probe set Ag7037, described in Table ZA. Results of the RTQ-PCR runs are shown in Tables ZB and ZC.

Table ZA. Probe Name Ag7037

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gctgttcagaaggctgtgaac-3'	21	554	336
Probe	TET-5'- acaggcaaaagactcatgcagtgtctcc -3'-TAMRA	28	583	337
Reverse	5'-ggccattttcacatcctctg-3'	20	617	338

Table ZB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag7037, Run 282263012	Tissue Name	Rel. Exp.(%) Ag7037, Run 282263012
AD 1 Hippo	0.0	Control (Path) 3 Temporal Ctx	0.0
AD 2 Hippo	7.5	Control (Path) 4 Temporal Ctx	37.9
AD 3 Hippo	3.7	AD 1 Occipital Ctx	6.8
AD 4 Hippo	0.0	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	63.3	AD 3 Occipital Ctx	0.0
AD 6 Hippo	46.7	AD 4 Occipital Ctx	0.0
Control 2 Hippo	0.0	AD 5 Occipital Ctx	4.1
Control 4 Hippo	2.8	AD 6 Occipital Ctx	18.4
Control (Path) 3 Hippo	13.0	Control 1 Occipital Ctx	4.4
AD 1 Temporal Ctx	21.6	Control 2 Occipital Ctx	29.9
AD 2 Temporal Ctx	11.7	Control 3 Occipital Ctx	11.0
AD 3 Temporal Ctx	0.0	Control 4 Occipital Ctx	16.2
AD 4 Temporal Ctx	21.3	Control (Path) 1 Occipital Ctx	40.3
AD 5 Inf Temporal Ctx	59.9	Control (Path) 2 Occipital Ctx	15.2
AD 5 Sup Temporal Ctx	28.9	Control (Path) 3 Occipital Ctx	0.0
AD 6 Inf Temporal Ctx	49.0	Control (Path) 4 Occipital Ctx	14.9
AD 6 Sup Temporal Ctx	45.4	Control 1 Parietal Ctx	16.7
Control 1 Temporal Ctx	0.0	Control 2 Parietal Ctx	35.8
Control 2 Temporal Ctx	35.6	Control 3 Parietal Ctx	12.2
Control 3 Temporal Ctx	24.7	Control (Path) 1 Parietal Ctx	37.6
Control 4 Temporal Ctx	4.3	Control (Path) 2 Parietal Ctx	4.9
Control (Path) 1 Temporal Ctx	6.9	Control (Path) 3 Parietal Ctx	0.0
Control (Path) 2 Temporal Ctx	100.0	Control (Path) 4 Parietal Ctx	14.4

Table ZC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag7037, Run 282263188	Tissue Name	Rel. Exp.(%) Ag7037, Run 282263188
Secondary Th1 act	0.0	HUVEC IL-1beta	0.1
Secondary Th2 act	0.0	HUVEC IFN gamma	0.3
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.5
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.1
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.3
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.2
Primary Th2 act	0.3	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL- 1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti- CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	100.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.2
Two Way MLR 3 day	0.0	HPAEC none	0.0

Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.5
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.2
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.1
B lymphocytes CD40L and IL-4	0.1	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.1
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	1.7
Monocytes rest	0.0	Neutrophils rest	0.2
Monocytes LPS	0.0	Colon	0.0
Macrophages rest	0.0	Lung	16.4
Macrophages LPS	0.0	Thymus	0.0
HUVEC none	0.0	Kidney	0.4
HUVEC starved	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag7037 This gene is expressed at low levels in the CNS. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurological disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

- 5 **Panel 4.1D Summary:** Ag7037 Highest expression of this gene is seen in liver cirrhosis (CT=29.6). Thus, expression of this gene could be used to differentiate between this sample and other samples on this panel and as a marker of this condition. Furthermore, therapeutic modulation of the expression or function of this gene may reduce or inhibit fibrosis that occurs in liver cirrhosis.

10 **AA. CG142092-03: C4b-binding protein alpha chain precursor protein-like protein.**

Expression of gene CG142092-03 was assessed using the primer-probe set Ag7668, described in Table AAA.

Table AAA. Probe Name Ag7668

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tgtgggtcctccaccact-3'	18	286	339
Probe	TET-5'- tctcagtcaacgtaatatccatcggggc a-3'-TAMRA	29	315	340
Reverse	5'- gttcaatttccagagtagttccagt-3'	25	355	341

CNS_neurodegeneration_v1.0 Summary: Ag7668 Expression of this gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4.1D Summary: Ag7668 Expression of this gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

AB. CG51117-03, CG51117-05, CG51117-06 and CG51117-07:

Nephronectin-like Protein

Expression of gene CG51117-03, CG51117-05, and CG51117-06 was assessed using the primer-probe sets Ag2505, Ag2667, Ag2767, Ag2831, Ag5113, Ag5124 and Ag7237, described in Tables ABA, ABB, ABC, ABD, ABE, ABF and ABG. Results of the RTQ-PCR runs are shown in Tables ABH, ABI, ABJ, ABK, ABL, ABM, ABN, ABO, ABP, ABQ, ABR and ABS. Note that Ag5113 is specific for CG51117-07 variant, and Ag5124 is specific for CG51117-06 variant.

Table ABA. Probe Name Ag2505

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-aaagaaggataccagggtagtg-3'	22	1113	342
Probe	TET-5'- atgattgaaccttcaggtccaattca -3'-TAMRA	26	1164	343
Reverse	5'-ggtaccatttccctttggtaca-3'	22	1190	344

Table ABB. Probe Name Ag2667

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gcagagaatagccaggataagg-3'	22	434	345
Probe	TET-5'- caaccacgatgcaaacatggtgaat-3'-TAMRA	25	477	346

Reverse	5' - cacttggttgcccgtac-3'	19	502	347
---------	---------------------------	----	-----	-----

Table ABC. Probe Name Ag2767

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - gcagagaatagccaggataagg-3'	22	434	348
Probe	TET-5' - caaccacgatgcaaactggtgaat-3' - TAMRA	25	477	349
Reverse	5' - cacttggttgcccgtac-3'	19	502	350

Table ABD. Probe Name Ag2831

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - gcagagaatagccaggataagg-3'	22	434	351
Probe	TET-5' - caaccacgatgcaaactggtgaat-3' - TAMRA	25	477	352
Reverse	5' - cacttggttgcccgtac-3'	19	502	353

Table ABE. Probe Name Ag5113

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - gtcagcctgtgtgccaa-3'	17	412	354
Probe	TET-5' - ccaaacaagtgcagtgtcatcctgg-3' - TAMRA	26	459	355
Reverse	5' - gggatgtgctcgtcttga-3'	18	506	356

5

Table ABF. Probe Name Ag5124

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - aggataaggtgccagctca-3'	19	447	357
Probe	TET-5' - ccaaacaagtgcagtgtcatcctgg-3' - TAMRA	26	510	358
Reverse	5' - gggatgtgctcgtcttga-3'	18	557	359

Table ABG. Probe Name Ag7237

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - gtgttcattccacggcaac-3'	19	1539	360

Probe	TET-5' - catcgtctgcactgactcctctttcta -3' -TAMRA	27	1588	361
Reverse	5' -gtgtaccagaacacctggatca- 3'	22	1625	362

Table ABH. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag2505, Run 248588456	Rel. Exp.(%) Ag2831, Run 244570250	Tissue Name	Rel. Exp.(%) Ag2505, Run 248588456	Rel. Exp.(%) Ag2831, Run 244570250
110967 COPD-F	15.3	9.3	112427 Match Control Psoriasis-F	16.7	18.9
110980 COPD-F	11.8	7.0	112418 Psoriasis-M	14.0	13.9
110968 COPD-M	8.9	5.8	112723 Match Control Psoriasis-M	0.2	0.2
110977 COPD-M	28.1	14.1	112419 Psoriasis-M	18.2	8.7
110989 Emphysema-F	9.6	12.2	112424 Match Control Psoriasis-M	6.8	6.7
110992 Emphysema-F	1.9	4.1	112420 Psoriasis-M	13.9	15.5
110993 Emphysema-F	7.7	9.3	112425 Match Control Psoriasis-M	13.6	16.6
110994 Emphysema-F	5.2	4.1	104689 (MF) OA Bone-Backus	25.3	38.4
110995 Emphysema-F	3.6	4.3	104690 (MF) Adj "Normal" Bone-Backus	27.9	21.2
110996 Emphysema-F	0.4	0.2	104691 (MF) OA Synovium-Backus	2.9	3.0
110997 Asthma-M	4.6	3.0	104692 (BA) OA Cartilage-Backus	0.0	0.0
111001 Asthma-F	2.3	5.0	104694 (BA) OA Bone-Backus	5.8	18.7
111002 Asthma-F	3.5	7.2	104695 (BA) Adj "Normal" Bone-Backus	14.1	19.8
111003 Atopic Asthma-F	22.5	22.2	104696 (BA) OA Synovium-Backus	2.3	3.6
111004 Atopic Asthma-F	10.4	11.4	104700 (SS) OA Bone-Backus	28.9	22.4

111005 Atopic Asthma-F	7.3	9.4	104701 (SS) Adj "Normal" Bone-Backus	25.5	18.7
111006 Atopic Asthma-F	1.6	1.4	104702 (SS) OA Synovium-Backus	11.7	7.1
111417 Allergy-M	5.1	2.8	117093 OA Cartilage Rep7	7.5	7.5
112347 Allergy-M	1.4	0.2	112672 OA Bone5	19.2	17.2
112349 Normal Lung-F	0.7	0.2	112673 OA Synovium5	6.4	3.8
112357 Normal Lung-F	7.1	6.4	112674 OA Synovial Fluid cells5	6.8	4.2
112354 Normal Lung-M	7.6	6.0	117100 OA Cartilage Rep14	1.9	2.1
112374 Crohns-F	9.0	3.2	112756 OA Bone9	26.4	31.6
112389 Match Control Crohns-F	11.2	6.6	112757 OA Synovium9	2.8	1.3
112375 Crohns-F	10.2	6.0	112758 OA Synovial Fluid Cells9	8.1	6.3
112732 Match Control Crohns-F	1.2	2.1	117125 RA Cartilage Rep2	14.8	9.2
112725 Crohns-M	0.9	1.6	113492 Bone2 RA	84.7	47.0
112387 Match Control Crohns-M	11.4	13.0	113493 Synovium2 RA	40.9	25.3
112378 Crohns-M	1.3	2.0	113494 Syn Fluid Cells RA	61.1	49.3
112390 Match Control Crohns-M	16.7	4.3	113499 Cartilage4 RA	90.1	73.2
112726 Crohns-M	21.8	17.0	113500 Bone4 RA	100.0	100.0
112731 Match Control Crohns-M	15.3	6.3	113501 Synovium4 RA	71.2	59.5
112380 Ulcer Col-F	5.8	7.0	113502 Syn Fluid Cells4 RA	48.6	37.9
112734 Match Control Ulcer Col-F	3.7	5.0	113495 Cartilage3 RA	77.9	47.0
112384 Ulcer Col-F	19.2	15.0	113496 Bone3 RA	92.0	41.8

112737 Match Control Ulcer Col-F	13.5	12.0	113497 Synovium3 RA	53.6	24.0
112386 Ulcer Col-F	8.4	6.0	113498 Syn Fluid Cells3 RA	98.6	57.0
112738 Match Control Ulcer Col-F	3.8	2.1	117106 Normal Cartilage Rep20	3.0	1.6
112381 Ulcer Col-M	5.0	9.9	113663 Bone3 Normal	2.0	0.7
112735 Match Control Ulcer Col-M	9.7	5.8	113664 Synovium3 Normal	0.5	0.4
112382 Ulcer Col-M	11.7	12.6	113665 Syn Fluid Cells3 Normal	1.6	1.4
112394 Match Control Ulcer Col-M	3.1	3.4	117107 Normal Cartilage Rep22	3.7	5.5
112383 Ulcer Col-M	5.1	13.7	113667 Bone4 Normal	4.3	6.0
112736 Match Control Ulcer Col-M	5.0	6.3	113668 Synovium4 Normal	9.1	7.4
112423 Psoriasis-F	14.0	10.3	113669 Syn Fluid Cells4 Normal	6.6	6.6

Table ABI. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag2505, Run 208123723	Rel. Exp.(%) Ag2505, Run 224116291	Rel. Exp.(%) Ag2667, Run 206955569	Rel. Exp.(%) Ag2767, Run 206985756	Rel. Exp.(%) Ag2831, Run 208699692	Rel. Exp.(%) Ag7237, Run 296423778
AD 1 Hippo	14.1	19.1	42.9	27.4	29.1	11.7
AD 2 Hippo	29.3	40.3	58.2	37.1	56.3	36.6
AD 3 Hippo	5.1	8.5	9.0	5.6	2.9	7.8
AD 4 Hippo	10.4	10.1	13.4	21.2	8.8	11.4
AD 5 Hippo	43.8	47.6	52.1	35.4	40.3	43.2
AD 6 Hippo	100.0	100.0	98.6	100.0	79.0	100.0
Control 2 Hippo	15.3	19.6	5.1	19.6	5.5	11.0
Control 4 Hippo	15.6	21.0	16.0	15.2	25.7	32.8
Control (Path) 3 Hippo	4.8	5.8	22.1	2.7	4.5	6.5
AD 1 Temporal Ctx	21.5	26.4	40.9	15.6	24.5	22.2

AD 2 Temporal Ctx	28.5	27.9	52.5	27.0	84.7	35.8
AD 3 Temporal Ctx	9.3	8.5	13.4	7.3	3.2	2.1
AD 4 Temporal Ctx	26.1	35.1	59.0	18.2	30.8	29.1
AD 5 Inf Temporal Ctx	28.9	33.9	39.5	27.0	49.3	37.1
AD 5 Sup Temporal Ctx	38.4	40.6	23.0	17.2	54.7	45.7
AD 6 Inf Temporal Ctx	83.5	96.6	100.0	66.4	100.0	94.0
AD 6 Sup Temporal Ctx	70.7	90.8	99.3	43.5	62.4	82.9
Control 1 Temporal Ctx	4.2	4.2	17.3	7.7	3.1	1.9
Control 2 Temporal Ctx	10.6	14.0	12.1	25.5	18.6	15.5
Control 3 Temporal Ctx	3.1	5.6	0.0	0.0	3.6	10.7
Control 3 Temporal Ctx	6.5	14.6	12.5	18.2	19.2	16.6
Control (Path) 1 Temporal Ctx	18.0	21.6	43.2	26.6	19.3	21.9
Control (Path) 2 Temporal Ctx	13.9	22.1	26.1	42.3	42.0	21.5
Control (Path) 3 Temporal Ctx	3.2	4.2	4.4	7.0	11.8	4.6
Control (Path) 4 Temporal Ctx	13.4	15.3	26.6	21.9	13.8	19.6
AD 1 Occipital Ctx	13.4	15.2	27.9	7.9	16.2	13.5
AD 2 Occipital Ctx (Missing)	0.0	0.0	0.0	0.0	0.0	0.0
AD 3 Occipital Ctx	4.1	5.8	11.8	0.0	9.9	3.0
AD 4 Occipital Ctx	19.3	23.2	17.0	9.0	41.2	27.4
AD 5 Occipital Ctx	17.8	16.8	39.5	27.9	17.6	19.6
AD 6 Occipital Ctx	29.3	43.8	30.8	25.5	13.2	32.8
Control 1 Occipital Ctx	4.0	3.0	14.0	0.0	13.0	2.8

Control 2 Occipital Ctx	21.8	25.3	2.3	29.1	17.7	32.3
Control 3 Occipital Ctx	6.9	7.3	28.7	4.8	7.1	9.2
Control 4 Occipital Ctx	9.4	10.3	13.8	9.1	17.6	17.7
Control (Path) 1 Occipital Ctx	29.1	28.1	37.6	29.3	47.6	34.2
Control (Path) 2 Occipital Ctx	5.1	7.0	6.7	5.3	8.6	5.3
Control (Path) 3 Occipital Ctx	1.6	2.5	25.7	3.5	0.0	2.4
Control (Path) 4 Occipital Ctx	13.7	17.2	19.8	13.5	13.7	12.2
Control 1 Parietal Ctx	3.8	4.0	10.8	24.1	7.3	3.8
Control 2 Parietal Ctx	37.4	47.6	53.6	36.1	57.4	53.6
Control 3 Parietal Ctx	4.1	5.4	0.0	3.4	3.5	5.3
Control (Path) 1 Parietal Ctx	23.5	28.9	24.7	21.5	42.6	30.4
Control (Path) 2 Parietal Ctx	15.7	20.2	44.8	11.5	39.5	20.6
Control (Path) 3 Parietal Ctx	2.6	4.0	14.9	3.7	3.0	2.4
Control (Path) 4 Parietal Ctx	21.9	25.7	49.7	20.4	50.7	26.2

Table ABJ. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag5113, Run 228738816	Rel. Exp.(%) Ag5124, Run 228745551	Tissue Name	Rel. Exp.(%) Ag5113, Run 228738816	Rel. Exp.(%) Ag5124, Run 228745551
Adipose	3.6	3.3	Renal ca. TK-10	0.0	0.0
Melanoma* Hs688(A).T	0.0	0.0	Bladder	1.2	2.4
Melanoma* Hs688(B).T	0.0	0.0	Gastric ca. (liver met.) NCI-N87	0.1	0.0
Melanoma* M14	0.0	0.0	Gastric ca. KATO III	0.1	0.0
Melanoma* LOXIMVI	0.0	0.0	Colon ca. SW-948	0.0	0.0
Melanoma* SK- MEL-5	0.0	0.0	Colon ca. SW480	0.0	0.0

Squamous cell carcinoma SCC-4	0.0	0.0	Colon ca.* (SW480 met) SW620	0.0	0.0
Testis Pool	3.9	5.9	Colon ca. HT29	0.0	0.0
Prostate ca.* (bone met) PC-3	0.1	0.0	Colon ca. HCT-116	0.1	0.5
Prostate Pool	7.3	5.0	Colon ca. CaCo-2	0.0	0.0
Placenta	0.1	0.0	Colon cancer tissue	0.5	1.3
Uterus Pool	4.7	4.1	Colon ca. SW1116	0.0	0.4
Ovarian ca. OVCAR-3	0.0	0.0	Colon ca. Colo-205	0.1	0.0
Ovarian ca. SK-OV-3	0.2	0.0	Colon ca. SW-48	0.0	0.0
Ovarian ca. OVCAR-4	0.1	0.0	Colon Pool	2.1	5.5
Ovarian ca. OVCAR-5	0.0	0.0	Small Intestine Pool	2.4	2.8
Ovarian ca. IGROV-1	0.0	1.4	Stomach Pool	2.0	3.2
Ovarian ca. OVCAR-8	0.3	0.0	Bone Marrow Pool	3.0	1.8
Ovary	0.5	2.5	Fetal Heart	0.2	0.0
Breast ca. MCF-7	0.2	0.6	Heart Pool	2.4	1.9
Breast ca. MDA-MB-231	0.0	0.0	Lymph Node Pool	8.5	9.3
Breast ca. BT 549	0.3	1.4	Fetal Skeletal Muscle	1.4	2.1
Breast ca. T47D	0.0	0.0	Skeletal Muscle Pool	1.1	1.7
Breast ca. MDA-N	0.0	0.0	Spleen Pool	0.3	0.0
Breast Pool	4.2	3.7	Thymus Pool	1.1	0.0
Trachea	6.0	4.2	CNS cancer (glio/astro) U87-MG	0.0	0.0
Lung	3.2	2.3	CNS cancer (glio/astro) U-118-MG	0.0	0.9
Fetal Lung	100.0	100.0	CNS cancer (neuro;met) SK-N-AS	0.0	0.0
Lung ca. NCI-N417	0.0	0.0	CNS cancer (astro) SF-539	0.3	0.5
Lung ca. LX-1	0.0	0.0	CNS cancer (astro) SNB-75	0.0	0.6
Lung ca. NCI-H146	0.0	0.0	CNS cancer (glio) SNB-19	0.2	1.0

Lung ca. SHP-77	0.0	0.0	CNS cancer (glio) SF-295	0.0	1.9
Lung ca. A549	0.0	0.0	Brain (Amygdala) Pool	0.0	0.0
Lung ca. NCI-H526	0.0	0.0	Brain (cerebellum)	0.0	0.0
Lung ca. NCI-H23	0.0	0.0	Brain (fetal)	0.2	0.5
Lung ca. NCI-H460	0.1	0.0	Brain (Hippocampus) Pool	0.5	0.8
Lung ca. HOP-62	0.0	0.0	Cerebral Cortex Pool	0.3	0.0
Lung ca. NCI-H522	0.0	0.0	Brain (Substantia nigra) Pool	0.1	0.0
Liver	0.0	0.0	Brain (Thalamus) Pool	0.0	0.0
Fetal Liver	0.2	0.0	Brain (whole)	0.3	0.5
Liver ca. HepG2	0.0	0.0	Spinal Cord Pool	0.0	0.4
Kidney Pool	13.2	8.2	Adrenal Gland	0.6	2.0
Fetal Kidney	0.3	1.5	Pituitary gland Pool	0.2	0.0
Renal ca. 786-0	0.0	0.0	Salivary Gland	0.0	0.0
Renal ca. A498	0.0	0.0	Thyroid (female)	0.3	0.4
Renal ca. ACHN	1.3	0.4	Pancreatic ca. CAPAN2	0.0	0.0
Renal ca. UO-31	0.0	0.0	Pancreas Pool	1.7	3.8

Table ABK. General_screening_panel_v1.6

Tissue Name	Rel. Exp.(%) Ag7237, Run 296433071	Tissue Name	Rel. Exp.(%) Ag7237, Run 296433071
Adipose	19.9	Renal ca. TK-10	2.4
Melanoma* Hs688(A).T	0.2	Bladder	19.2
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	45.7
Melanoma* M14	0.0	Gastric ca. KATO III	11.7
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	9.1
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.5
Squamous cell carcinoma SCC-4	1.9	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	5.0	Colon ca. HT29	12.4
Prostate ca.* (bone met) PC-3	0.3	Colon ca. HCT-116	10.0
Prostate Pool	44.1	Colon ca. CaCo-2	17.4
Placenta	1.0	Colon cancer tissue	5.9
Uterus Pool	2.0	Colon ca. SW1116	4.8
Ovarian ca. OVCAR-3	5.6	Colon ca. Colo-205	4.2

Ovarian ca. SK-OV-3	18.0	Colon ca. SW-48	10.1
Ovarian ca. OVCAR-4	0.0	Colon Pool	5.8
Ovarian ca. OVCAR-5	4.7	Small Intestine Pool	6.3
Ovarian ca. IGROV-1	10.9	Stomach Pool	4.3
Ovarian ca. OVCAR-8	0.9	Bone Marrow Pool	6.4
Ovary	2.3	Fetal Heart	6.0
Breast ca. MCF-7	71.7	Heart Pool	4.5
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	18.0
Breast ca. BT 549	12.2	Fetal Skeletal Muscle	12.8
Breast ca. T47D	6.2	Skeletal Muscle Pool	0.6
Breast ca. MDA-N	0.0	Spleen Pool	5.7
Breast Pool	7.5	Thymus Pool	6.9
Trachea	10.7	CNS cancer (glio/astro) U87-MG	0.0
Lung	2.4	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	100.0	CNS cancer (neuro;met) SK-N-AS	0.1
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF- 539	1.7
Lung ca. LX-1	5.2	CNS cancer (astro) SNB- 75	0.7
Lung ca. NCI-H146	7.9	CNS cancer (glio) SNB- 19	12.7
Lung ca. SHP-77	0.7	CNS cancer (glio) SF-295	0.2
Lung ca. A549	0.7	Brain (Amygdala) Pool	3.0
Lung ca. NCI-H526	0.3	Brain (cerebellum)	0.6
Lung ca. NCI-H23	4.5	Brain (fetal)	24.8
Lung ca. NCI-H460	0.5	Brain (Hippocampus) Pool	7.5
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	3.7
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	1.9
Liver	0.0	Brain (Thalamus) Pool	5.5
Fetal Liver	1.3	Brain (whole)	6.1
Liver ca. HepG2	2.7	Spinal Cord Pool	1.0
Kidney Pool	0.0	Adrenal Gland	3.1
Fetal Kidney	23.2	Pituitary gland Pool	5.8
Renal ca. 786-0	0.0	Salivary Gland	0.7
Renal ca. A498	0.0	Thyroid (female)	47.3
Renal ca. ACHN	47.6	Pancreatic ca. CAPAN2	0.7
Renal ca. UO-31	0.0	Pancreas Pool	9.6

Table ABL. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag2505, Run 165531061	Rel. Exp.(%) Ag2667, Run 162554578	Rel. Exp.(%) Ag2767, Run 165527179	Rel. Exp.(%) Ag2831, Run 165517578
Liver adenocarcinoma	1.8	0.0	0.0	1.3
Pancreas	13.7	8.9	35.4	14.1
Pancreatic ca. CAPAN 2	1.5	2.0	0.0	2.0
Adrenal gland	3.6	2.9	2.8	4.6
Thyroid	100.0	52.5	100.0	67.8
Salivary gland	4.1	2.3	9.3	2.5
Pituitary gland	37.6	9.9	18.7	19.8
Brain (fetal)	44.1	6.8	40.9	28.5
Brain (whole)	9.3	0.6	11.2	0.0
Brain (amygdala)	8.1	4.4	8.2	2.6
Brain (cerebellum)	1.8	0.8	0.0	4.4
Brain (hippocampus)	10.2	1.6	6.4	2.2
Brain (substantia nigra)	29.3	3.7	12.5	11.0
Brain (thalamus)	3.6	1.9	8.7	7.2
Cerebral Cortex	7.7	8.8	3.8	1.3
Spinal cord	15.2	14.4	13.3	10.4
glio/astro U87-MG	0.0	0.0	0.0	0.0
glio/astro U-118-MG	0.0	0.0	0.0	0.0
astrocytoma SW1783	0.3	0.6	0.0	1.0
neuro*; met SK-N-AS	0.4	0.0	0.0	0.0
astrocytoma SF-539	1.8	1.2	2.5	1.2
astrocytoma SNB-75	2.7	0.6	0.0	2.0
glioma SNB-19	0.0	0.0	0.0	0.0
glioma U251	9.3	2.0	12.2	9.1
glioma SF-295	0.4	0.0	2.6	1.3
Heart (fetal)	10.0	24.7	9.6	7.4
Heart	3.1	0.0	2.4	2.4
Skeletal muscle (fetal)	12.8	66.4	7.7	1.3
Skeletal muscle	20.9	2.1	7.5	13.1
Bone marrow	1.2	1.9	4.5	0.9
Thymus	6.0	24.0	17.8	4.9
Spleen	6.7	5.0	19.8	9.2
Lymph node	6.7	1.4	5.4	2.6
Colorectal	23.5	19.3	6.5	9.9
Stomach	12.0	1.8	4.5	2.5
Small intestine	54.3	13.3	69.7	43.2
Colon ca. SW480	1.1	0.5	0.0	0.0
Colon ca.* SW620(SW480 met)	1.4	0.0	2.8	3.1
Colon ca. HT29	7.3	28.3	11.3	5.3

Colon ca. HCT-116	7.3	9.0	12.4	10.2
Colon ca. CaCo-2	10.7	29.7	19.6	11.1
Colon ca. tissue(ODO3866)	8.5	14.6	10.6	10.6
Colon ca. HCC-2998	2.9	6.3	19.8	2.9
Gastric ca.* (liver met) NCI-N87	71.7	49.7	95.3	100.0
Bladder	14.8	44.4	29.7	20.4
Trachea	21.9	13.9	18.6	5.5
Kidney	38.2	74.2	56.3	67.4
Kidney (fetal)	27.5	37.1	40.1	33.9
Renal ca. 786-0	0.0	0.0	0.0	0.0
Renal ca. A498	0.2	1.7	3.7	2.9
Renal ca. RXF 393	39.8	5.9	20.3	10.8
Renal ca. ACHN	51.1	6.8	20.2	7.2
Renal ca. UO-31	0.2	0.0	0.0	0.0
Renal ca. TK-10	0.0	0.0	0.0	0.0
Liver	1.4	0.7	0.0	3.5
Liver (fetal)	2.3	0.0	4.0	1.2
Liver ca. (hepatoblast) HepG2	11.8	4.7	19.5	10.8
Lung	75.3	46.0	91.4	84.1
Lung (fetal)	54.7	100.0	92.0	64.6
Lung ca. (small cell) LX-1	5.5	2.5	5.8	4.1
Lung ca. (small cell) NCI-H69	5.6	6.2	10.1	5.3
Lung ca. (s.cell var.) SHP-77	0.2	0.6	0.0	0.0
Lung ca. (large cell) NCI-H460	1.2	0.0	0.0	0.0
Lung ca. (non-sm. cell) A549	1.2	0.7	3.1	1.3
Lung ca. (non-s.cell) NCI-H23	3.2	4.7	8.2	4.5
Lung ca. (non-s.cell) HOP-62	0.0	0.0	0.0	0.0
Lung ca. (non-s.cl) NCI-H522	0.0	0.0	0.0	0.0
Lung ca. (squam.) SW 900	1.8	0.0	3.4	2.0
Lung ca. (squam.) NCI-H596	14.6	10.9	31.6	53.6
Mammary gland	11.9	4.9	23.8	17.4
Breast ca.* (pl.ef) MCF-7	89.5	92.7	84.1	80.1

Breast ca.* (pl.ef) MDA-MB-231	0.0	0.0	0.0	0.0
Breast ca.* (pl.ef) T47D	24.7	7.6	20.3	9.9
Breast ca. BT-549	2.3	0.0	0.0	1.2
Breast ca. MDA-N	0.0	0.0	0.0	0.0
Ovary	3.5	20.3	8.5	4.9
Ovarian ca. OVCAR-3	6.4	2.6	8.2	4.5
Ovarian ca. OVCAR-4	0.0	0.0	0.0	0.0
Ovarian ca. OVCAR-5	0.0	0.0	0.0	0.0
Ovarian ca. OVCAR-8	0.2	1.0	0.0	0.0
Ovarian ca. IGROV-1	14.5	17.6	31.6	33.7
Ovarian ca.* (ascites) SK-OV-3	9.3	5.4	20.7	22.5
Uterus	27.7	3.5	39.0	46.3
Placenta	2.9	4.9	6.4	8.9
Prostate	25.0	8.8	16.7	16.7
Prostate ca.* (bone met)PC-3	0.0	0.0	0.0	0.0
Testis	2.5	0.7	4.4	2.7
Melanoma Hs688(A).T	0.0	0.0	0.0	0.0
Melanoma* (met) Hs688(B).T	0.0	0.0	0.0	0.0
Melanoma UACC-62	0.0	0.0	0.0	0.0
Melanoma M14	0.0	0.0	0.0	0.0
Melanoma LOX IMVI	0.0	0.0	0.0	0.0
Melanoma* (met) SK- MEL-5	0.0	0.0	0.0	0.0
Adipose	19.3	6.5	4.3	22.1

Table ABM. Panel 2.2

Tissue Name	Rel. Exp.(%) Ag2831, Run 175063921	Tissue Name	Rel. Exp.(%) Ag2831, Run 175063921
Normal Colon	4.7	Kidney Margin (OD04348)	100.0
Colon cancer (OD06064)	24.7	Kidney malignant cancer (OD06204B)	0.0
Colon Margin (OD06064)	12.0	Kidney normal adjacent tissue (OD06204E)	7.0
Colon cancer (OD06159)	1.1	Kidney Cancer (OD04450-01)	1.2
Colon Margin (OD06159)	6.2	Kidney Margin (OD04450-03)	16.6
Colon cancer (OD06297- 04)	1.9	Kidney Cancer 8120613	1.8

Colon Margin (OD06297-05)	6.9	Kidney Margin 8120614	5.7
CC Gr.2 ascend colon (ODO3921)	0.4	Kidney Cancer 9010320	0.6
CC Margin (ODO3921)	2.7	Kidney Margin 9010321	2.6
Colon cancer metastasis (OD06104)	2.4	Kidney Cancer 8120607	6.2
Lung Margin (OD06104)	10.2	Kidney Margin 8120608	2.3
Colon mets to lung (OD04451-01)	7.0	Normal Uterus	13.4
Lung Margin (OD04451-02)	20.4	Uterine Cancer 064011	0.8
Normal Prostate	4.9	Normal Thyroid	6.1
Prostate Cancer (OD04410)	5.9	Thyroid Cancer 064010	28.5
Prostate Margin (OD04410)	8.3	Thyroid Cancer A302152	46.3
Normal Ovary	1.9	Thyroid Margin A302153	21.0
Ovarian cancer (OD06283-03)	1.2	Normal Breast	10.2
Ovarian Margin (OD06283-07)	3.6	Breast Cancer (OD04566)	1.5
Ovarian Cancer 064008	7.8	Breast Cancer 1024	4.6
Ovarian cancer (OD06145)	0.9	Breast Cancer (OD04590-01)	62.0
Ovarian Margin (OD06145)	0.9	Breast Cancer Mets (OD04590-03)	98.6
Ovarian cancer (OD06455-03)	0.0	Breast Cancer Metastasis (OD04655-05)	70.7
Ovarian Margin (OD06455-07)	7.3	Breast Cancer 064006	3.6
Normal Lung	14.2	Breast Cancer 9100266	3.4
Invasive poor diff. lung adeno (ODO4945-01)	1.5	Breast Margin 9100265	2.9
Lung Margin (ODO4945-03)	15.5	Breast Cancer A209073	1.7
Lung Malignant Cancer (OD03126)	4.2	Breast Margin A2090734	2.5
Lung Margin (OD03126)	8.3	Breast cancer (OD06083)	49.7
Lung Cancer (OD05014A)	5.4	Breast cancer node metastasis (OD06083)	64.2
Lung Margin (OD05014B)	41.5	Normal Liver	0.5
Lung cancer (OD06081)	3.8	Liver Cancer 1026	0.5
Lung Margin (OD06081)	37.6	Liver Cancer 1025	1.8
Lung Cancer (OD04237-01)	1.6	Liver Cancer 6004-T	0.0
Lung Margin (OD04237-02)	33.2	Liver Tissue 6004-N	1.3

Ocular Melanoma Metastasis	0.0	Liver Cancer 6005-T	0.5
Ocular Melanoma Margin (Liver)	0.0	Liver Tissue 6005-N	1.4
Melanoma Metastasis	0.0	Liver Cancer 064003	0.0
Melanoma Margin (Lung)	37.9	Normal Bladder	2.8
Normal Kidney	5.5	Bladder Cancer 1023	2.5
Kidney Ca, Nuclear grade 2 (OD04338)	26.6	Bladder Cancer A302173	6.2
Kidney Margin (OD04338)	0.9	Normal Stomach	2.8
Kidney Ca Nuclear grade 1/2 (OD04339)	2.0	Gastric Cancer 9060397	0.0
Kidney Margin (OD04339)	10.3	Stomach Margin 9060396	1.4
Kidney Ca, Clear cell type (OD04340)	4.5	Gastric Cancer 9060395	2.3
Kidney Margin (OD04340)	12.6	Stomach Margin 9060394	4.1
Kidney Ca, Nuclear grade 3 (OD04348)	1.4	Gastric Cancer 064005	5.7

Table ABN. Panel 2D

Tissue Name	Rel. Exp.(%) Ag2667, Run 162558279	Rel. Exp.(%) Ag2767, Run 162555855	Rel. Exp.(%) Ag2831, Run 163578438	Tissue Name	Rel. Exp.(%) Ag2667, Run 162558279	Rel. Exp.(%) Ag2767, Run 162555855	Rel. Exp.(%) Ag2831, Run 163578438
Normal Colon	4.8	4.8	6.1	Kidney Margin 8120608	2.3	2.4	2.0
CC Well to Mod Diff (ODO3866)	1.1	0.8	0.9	Kidney Cancer 8120613	8.5	9.9	9.6
CC Margin (ODO3866)	0.8	1.2	1.5	Kidney Margin 8120614	3.0	3.2	2.8
CC Gr.2 rectosigmoid (ODO3868)	0.8	0.5	0.3	Kidney Cancer 9010320	1.0	1.6	0.9
CC Margin (ODO3868)	0.2	0.2	0.1	Kidney Margin 9010321	3.9	4.6	0.0
CC Mod Diff (ODO3920)	0.2	0.2	0.1	Normal Uterus	1.3	1.1	0.6
CC Margin (ODO3920)	1.0	0.7	0.9	Uterus Cancer 064011	1.6	1.1	1.4

CC Gr.2 ascend colon (ODO3921)	3.8	3.8	3.8	Normal Thyroid	13.9	13.7	10.4
CC Margin (ODO3921)	1.4	1.5	1.0	Thyroid Cancer 064010	33.2	35.1	36.9
CC from Partial Hepatectomy (ODO4309) Mets	6.5	5.6	6.1	Thyroid Cancer A302152	19.3	21.3	21.5
Liver Margin (ODO4309)	0.3	0.4	0.1	Thyroid Margin A302153	41.8	39.8	37.9
Colon mets to lung (OD04451-01)	1.5	1.6	1.2	Normal Breast	1.7	2.4	1.7
Lung Margin (OD04451-02)	3.1	4.0	3.3	Breast Cancer (OD04566)	2.0	2.2	2.5
Normal Prostate 6546-1	10.6	10.5	11.9	Breast Cancer (OD04590- 01)	68.3	69.7	64.2
Prostate Cancer (OD04410)	13.3	13.4	15.2	Breast Cancer Mets (OD04590- 03)	100.0	100.0	100.0
Prostate Margin (OD04410)	8.3	12.1	10.5	Breast Cancer Metastasis (OD04655- 05)	38.7	39.0	33.4
Prostate Cancer (OD04720-01)	2.2	1.7	1.6	Breast Cancer 064006	3.7	4.0	3.9
Prostate Margin (OD04720-02)	5.0	4.5	4.1	Breast Cancer 1024	2.7	2.1	2.7
Normal Lung 061010	15.0	17.3	13.5	Breast Cancer 9100266	2.6	2.3	2.9
Lung Met to Muscle (ODO4286)	0.5	0.5	0.3	Breast Margin 9100265	0.9	0.8	0.6
Muscle Margin (ODO4286)	0.1	0.1	0.0	Breast Cancer A209073	3.5	3.7	3.8
Lung Malignant Cancer (OD03126)	3.7	3.8	3.5	Breast Margin A209073	1.7	1.5	1.5
Lung Margin (OD03126)	15.1	20.4	15.5	Normal Liver	0.2	0.1	0.1

Lung Cancer (OD04404)	4.7	4.2	2.8	Liver Cancer 064003	0.0	0.1	0.2
Lung Margin (OD04404)	12.1	12.9	9.8	Liver Cancer 1025	0.1	0.1	0.0
Lung Cancer (OD04565)	0.6	0.7	0.4	Liver Cancer 1026	0.8	0.5	0.8
Lung Margin (OD04565)	9.9	8.4	8.6	Liver Cancer 6004-T	0.1	0.1	0.1
Lung Cancer (OD04237-01)	1.1	1.5	1.0	Liver Tissue 6004-N	0.8	0.9	0.9
Lung Margin (OD04237-02)	17.4	13.0	14.3	Liver Cancer 6005-T	0.4	0.7	0.5
Ocular Mel Met to Liver (ODO4310)	0.0	0.1	0.0	Liver Tissue 6005-N	0.1	0.1	0.1
Liver Margin (ODO4310)	0.2	0.2	0.3	Normal Bladder	3.5	3.8	4.2
Melanoma Mets to Lung (OD04321)	0.1	0.3	0.2	Bladder Cancer 1023	0.9	0.9	0.6
Lung Margin (OD04321)	21.3	20.7	19.5	Bladder Cancer A302173	3.8	4.6	4.4
Normal Kidney	14.9	18.4	15.2	Bladder Cancer (OD04718-01)	0.6	1.1	1.0
Kidney Ca, Nuclear grade 2 (OD04338)	0.9	1.2	0.6	Bladder Normal Adjacent (OD04718-03)	0.6	0.3	0.8
Kidney Margin (OD04338)	7.3	10.3	7.0	Normal Ovary	0.8	0.7	0.6
Kidney Ca Nuclear grade 1/2 (OD04339)	0.3	0.3	0.6	Ovarian Cancer 064008	7.2	9.6	8.8
Kidney Margin (OD04339)	14.8	11.7	14.6	Ovarian Cancer (OD04768-07)	0.2	0.2	0.1
Kidney Ca, Clear cell type (OD04340)	6.5	7.8	8.1	Ovary Margin (OD04768-08)	1.5	1.8	1.3
Kidney Margin (OD04340)	11.0	9.2	9.8	Normal Stomach	0.5	1.0	0.6
Kidney Ca, Nuclear grade 3 (OD04348)	1.1	0.6	1.1	Gastric Cancer 9060358	0.3	0.2	0.6

Kidney Margin (OD04348)	15.5	11.7	13.5	Stomach Margin 9060359	0.4	0.9	0.8
Kidney Cancer (OD04622-01)	1.7	1.2	1.6	Gastric Cancer 9060395	1.6	2.2	1.3
Kidney Margin (OD04622-03)	3.0	2.7	2.5	Stomach Margin 9060394	0.7	1.2	0.6
Kidney Cancer (OD04450-01)	0.1	0.2	0.3	Gastric Cancer 9060397	0.3	0.3	0.0
Kidney Margin (OD04450-03)	11.6	15.2	14.0	Stomach Margin 9060396	0.3	0.3	0.0
Kidney Cancer 8120607	2.6	2.6	2.9	Gastric Cancer 064005	11.7	16.5	11.6

Table ABO. Panel 3D

Tissue Name	Rel. Exp.(%) Ag2831, Run 164843468	Tissue Name	Rel. Exp.(%) Ag2831, Run 164843468
Daoy- Medulloblastoma	0.4	Ca Ski- Cervical epidermoid carcinoma (metastasis)	9.4
TE671- Medulloblastoma	3.7	ES-2- Ovarian clear cell carcinoma	0.0
D283 Med- Medulloblastoma	6.7	Ramos- Stimulated with PMA/ionomycin 6h	0.0
PFSK-1- Primitive Neuroectodermal	0.0	Ramos- Stimulated with PMA/ionomycin 14h	0.0
XF-498- CNS	1.2	MEG-01- Chronic myelogenous leukemia (megakaryoblast)	0.7
SNB-78- Glioma	1.7	Raji- Burkitt's lymphoma	0.0
SF-268- Glioblastoma	0.0	Daudi- Burkitt's lymphoma	0.0
T98G- Glioblastoma	0.0	U266- B-cell plasmacytoma	0.0
SK-N-SH- Neuroblastoma (metastasis)	0.0	CA46- Burkitt's lymphoma	0.0
SF-295- Glioblastoma	0.0	RL- non-Hodgkin's B-cell lymphoma	0.0
Cerebellum	0.0	JM1- pre-B-cell lymphoma	0.0
Cerebellum	0.0	Jurkat- T cell leukemia	0.0
NCI-H292- Mucoepidermoid lung carcinoma	23.7	TF-1- Erythroleukemia	0.0

DMS-114- Small cell lung cancer	0.0	HUT 78- T-cell lymphoma	0.0
DMS-79- Small cell lung cancer	1.1	U937- Histiocytic lymphoma	0.0
NCI-H146- Small cell lung cancer	100.0	KU-812- Myelogenous leukemia	0.6
NCI-H526- Small cell lung cancer	5.6	769-P- Clear cell renal carcinoma	3.2
NCI-N417- Small cell lung cancer	0.8	Caki-2- Clear cell renal carcinoma	0.8
NCI-H82- Small cell lung cancer	0.0	SW 839- Clear cell renal carcinoma	0.9
NCI-H157- Squamous cell lung cancer (metastasis)	0.0	G401- Wilms' tumor	0.0
NCI-H1155- Large cell lung cancer	14.6	Hs766T- Pancreatic carcinoma (LN metastasis)	0.0
NCI-H1299- Large cell lung cancer	0.0	CAPAN-1- Pancreatic adenocarcinoma (liver metastasis)	0.0
NCI-H727- Lung carcinoid	14.8	SU86.86- Pancreatic carcinoma (liver metastasis)	0.0
NCI-UMC-11- Lung carcinoid	84.1	BxPC-3- Pancreatic adenocarcinoma	0.0
LX-1- Small cell lung cancer	7.5	HPAC- Pancreatic adenocarcinoma	0.0
Colo-205- Colon cancer	18.7	MIA PaCa-2- Pancreatic carcinoma	0.0
KM12- Colon cancer	66.4	CFPAC-1- Pancreatic ductal adenocarcinoma	0.0
KM20L2- Colon cancer	8.4	PANC-1- Pancreatic epithelioid ductal carcinoma	0.6
NCI-H716- Colon cancer	23.2	T24- Bladder carcinoma (transitional cell)	0.0
SW-48- Colon adenocarcinoma	63.7	5637- Bladder carcinoma	0.0
SW1116- Colon adenocarcinoma	15.5	HT-1197- Bladder carcinoma	3.2
LS 174T- Colon adenocarcinoma	62.9	UM-UC-3- Bladder carcinoma (transitional cell)	0.0
SW-948- Colon adenocarcinoma	2.7	A204- Rhabdomyosarcoma	0.0
SW-480- Colon adenocarcinoma	39.2	HT-1080- Fibrosarcoma	0.0
NCI-SNU-5- Gastric carcinoma	0.0	MG-63- Osteosarcoma	0.0
KATO III- Gastric carcinoma	33.0	SK-LMS-1- Leiomyosarcoma (vulva)	0.0

NCI-SNU-16- Gastric carcinoma	0.0	SJRH30- Rhabdomyosarcoma (met to bone marrow)	27.7
NCI-SNU-1- Gastric carcinoma	35.6	A431- Epidermoid carcinoma	17.8
RF-1- Gastric adenocarcinoma	0.0	WM266-4- Melanoma	0.0
RF-48- Gastric adenocarcinoma	0.7	DU 145- Prostate carcinoma (brain metastasis)	0.0
MKN-45- Gastric carcinoma	96.6	MDA-MB-468- Breast adenocarcinoma	2.7
NCI-N87- Gastric carcinoma	79.6	SCC-4- Squamous cell carcinoma of tongue	0.0
OVCAR-5- Ovarian carcinoma	0.0	SCC-9- Squamous cell carcinoma of tongue	0.0
RL95-2- Uterine carcinoma	0.0	SCC-15- Squamous cell carcinoma of tongue	0.0
HelaS3- Cervical adenocarcinoma	0.0	CAL 27- Squamous cell carcinoma of tongue	1.2

Table ABP. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag2831, Run 244570230	Rel. Exp.(%) Ag5124, Run 225784387	Tissue Name	Rel. Exp.(%) Ag2831, Run 244570230	Rel. Exp.(%) Ag5124, Run 225784387
Secondary Th1 act	0.0	0.0	HUVEC IL-1beta	0.0	0.0
Secondary Th2 act	0.0	0.0	HUVEC IFN gamma	0.0	9.0
Secondary Tr1 act	0.0	0.0	HUVEC TNF alpha + IFN gamma	0.0	0.0
Secondary Th1 rest	0.0	0.0	HUVEC TNF alpha + IL4	0.0	0.0
Secondary Th2 rest	0.0	0.0	HUVEC IL-11	0.0	0.0
Secondary Tr1 rest	0.0	0.0	Lung Microvascular EC none	0.0	0.0
Primary Th1 act	0.0	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0	0.0
Primary Th2 act	0.0	0.0	Microvascular Dermal EC none	0.0	0.0
Primary Tr1 act	0.0	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0	0.0
Primary Th1 rest	0.0	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0	0.0
Primary Th2 rest	0.0	0.0	Small airway epithelium none	0.0	0.0

Primary Tr1 rest	0.0	0.0	Small airway epithelium TNFalpha + IL-1beta	11.5	0.0
CD45RA CD4 lymphocyte act	0.0	0.0	Coronary artery SMC rest	0.0	0.0
CD45RO CD4 lymphocyte act	0.0	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0	0.0
CD8 lymphocyte act	0.0	0.0	Astrocytes rest	0.9	0.0
Secondary CD8 lymphocyte rest	0.0	0.0	Astrocytes TNFalpha + IL-1beta	9.9	0.0
Secondary CD8 lymphocyte act	0.0	0.0	KU-812 (Basophil) rest	0.0	0.0
CD4 lymphocyte none	0.0	0.0	KU-812 (Basophil) PMA/ionomycin	0.0	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	0.0	CCD1106 (Keratinocytes) none	0.8	0.0
LAK cells rest	0.0	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	4.5	0.0
LAK cells IL-2	0.0	0.0	Liver cirrhosis	14.9	19.9
LAK cells IL-2+IL-12	0.0	0.0	NCI-H292 none	19.9	0.0
LAK cells IL-2+IFN gamma	0.0	0.0	NCI-H292 IL-4	62.4	0.0
LAK cells IL-2+ IL-18	0.0	0.0	NCI-H292 IL-9	57.8	0.0
LAK cells PMA/ionomycin	0.0	0.0	NCI-H292 IL-13	73.2	0.0
NK Cells IL-2 rest	0.0	0.0	NCI-H292 IFN gamma	21.0	0.0
Two Way MLR 3 day	0.0	0.0	HPAEC none	0.0	0.0
Two Way MLR 5 day	0.0	0.0	HPAEC TNF alpha + IL-1 beta	3.6	0.0
Two Way MLR 7 day	0.0	0.0	Lung fibroblast none	17.0	0.0
PBMC rest	0.0	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0	0.0
PBMC PWM	0.0	0.0	Lung fibroblast IL-4	9.7	48.3
PBMC PHA-L	0.0	0.0	Lung fibroblast IL-9	6.7	0.0
Ramos (B cell) none	0.0	0.0	Lung fibroblast IL-13	1.6	17.1
Ramos (B cell) ionomycin	0.0	0.0	Lung fibroblast IFN gamma	21.3	26.2
B lymphocytes PWM	0.0	0.0	Dermal fibroblast CCD1070 rest	0.0	0.0
B lymphocytes CD40L and IL-4	1.3	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0	0.0

EOL-1 dbcAMP	0.0	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	0.0	Dermal fibroblast IFN gamma	0.0	0.0
Dendritic cells none	0.0	0.0	Dermal fibroblast IL-4	0.0	0.0
Dendritic cells LPS	0.0	0.0	Dermal Fibroblasts rest	0.0	0.0
Dendritic cells anti- CD40	0.0	0.0	Neutrophils TNFa+LPS	0.0	0.0
Monocytes rest	0.0	0.0	Neutrophils rest	0.0	0.0
Monocytes LPS	0.0	0.0	Colon	2.2	8.7
Macrophages rest	0.0	0.0	Lung	2.1	100.0
Macrophages LPS	0.0	0.0	Thymus	1.8	0.0
HUVEC none	0.0	0.0	Kidney	100.0	36.9
HUVEC starved	0.0	0.0			

Table ABQ. Panel 4D

Tissue Name	Rel. Exp.(%) Ag2505, Run 164318134	Rel. Exp.(%) Ag2667, Run 158912431	Rel. Exp.(%) Ag2767, Run 162015289	Rel. Exp.(%) Ag2831, Run 162350949
Secondary Th1 act	0.0	0.0	0.0	0.0
Secondary Th2 act	0.0	0.0	0.0	0.0
Secondary Tr1 act	0.2	0.0	0.0	0.0
Secondary Th1 rest	0.0	0.0	0.0	0.0
Secondary Th2 rest	0.3	0.0	0.0	0.0
Secondary Tr1 rest	0.0	0.0	0.0	0.4
Primary Th1 act	0.0	0.0	0.0	0.0
Primary Th2 act	0.2	0.0	0.0	0.0
Primary Tr1 act	0.1	0.0	0.0	0.0
Primary Th1 rest	0.1	0.0	0.0	0.4
Primary Th2 rest	0.0	0.0	0.0	0.0
Primary Tr1 rest	0.1	0.0	0.3	0.0
CD45RA CD4 lymphocyte act	0.0	0.0	0.0	0.0
CD45RO CD4 lymphocyte act	0.0	0.0	0.0	0.0
CD8 lymphocyte act	0.0	0.0	0.0	0.0
Secondary CD8 lymphocyte rest	0.0	0.0	0.0	0.0
Secondary CD8 lymphocyte act	0.0	0.0	0.0	0.0
CD4 lymphocyte none	0.0	0.0	0.0	0.0
2ry Th1/Th2/Tr1_anti- CD95 CH11	0.0	0.0	0.0	0.0

LAK cells rest	0.0	0.0	0.0	0.0
LAK cells IL-2	0.0	0.0	0.0	0.0
LAK cells IL-2+IL-12	0.0	0.0	0.0	0.0
LAK cells IL-2+IFN gamma	0.0	0.0	0.0	0.0
LAK cells IL-2+ IL-18	0.0	0.0	0.0	0.0
LAK cells PMA/ionomycin	0.0	0.0	0.0	0.0
NK Cells IL-2 rest	0.0	0.0	0.0	0.0
Two Way MLR 3 day	0.0	0.0	0.0	0.0
Two Way MLR 5 day	0.0	0.0	0.0	0.0
Two Way MLR 7 day	0.0	0.0	0.0	0.0
PBMC rest	0.0	0.0	0.0	0.0
PBMC PWM	0.1	0.0	0.0	0.0
PBMC PHA-L	0.0	0.6	0.0	0.0
Ramos (B cell) none	0.0	0.0	0.0	0.0
Ramos (B cell) ionomycin	0.0	0.3	0.0	0.0
B lymphocytes PWM	0.2	0.3	0.9	0.7
B lymphocytes CD40L and IL-4	0.1	0.7	0.0	0.0
EOL-1 dbcAMP	0.0	0.0	0.0	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	0.0	0.0	0.0
Dendritic cells none	0.1	0.0	0.0	0.3
Dendritic cells LPS	0.0	0.0	0.0	0.0
Dendritic cells anti-CD40	0.0	0.0	0.0	0.0
Monocytes rest	0.0	0.0	0.0	0.0
Monocytes LPS	0.0	0.0	0.0	0.0
Macrophages rest	0.0	0.0	0.3	0.0
Macrophages LPS	0.0	0.0	0.0	0.0
HUVEC none	0.0	0.0	0.0	0.0
HUVEC starved	0.3	0.0	0.0	0.0
HUVEC IL-1beta	0.2	0.3	0.0	0.0
HUVEC IFN gamma	0.1	0.0	0.0	0.0
HUVEC TNF alpha + IFN gamma	0.0	0.0	0.0	0.7
HUVEC TNF alpha + IL4	0.1	0.0	0.0	0.0
HUVEC IL-11	0.0	0.0	0.0	0.0
Lung Microvascular EC none	0.0	0.0	0.0	0.0
Lung Microvascular EC TNFalpha + IL-1beta	0.0	0.3	0.0	0.0
Microvascular Dermal EC none	0.0	0.0	0.0	0.0

Microvascular Dermal EC TNFalpha + IL-1beta	0.2	0.0	1.4	0.0
Bronchial epithelium TNFalpha + IL1beta	2.0	0.0	1.2	0.0
Small airway epithelium none	0.5	0.0	0.0	0.4
Small airway epithelium TNFalpha + IL-1beta	19.6	10.4	11.7	8.4
Coronary artery SMC rest	0.0	0.0	0.0	0.0
Coronary artery SMC TNFalpha + IL-1beta	0.0	0.0	0.0	0.0
Astrocytes rest	2.0	2.6	1.9	2.3
Astrocytes TNFalpha + IL-1beta	2.0	4.6	8.2	4.4
KU-812 (Basophil) rest	0.0	0.0	0.0	0.0
KU-812 (Basophil) PMA/ionomycin	0.0	0.3	0.0	0.0
CCD1106 (Keratinocytes) none	0.4	0.4	0.8	0.6
CCD1106 (Keratinocytes) TNFalpha + IL-1beta	1.3	0.0	2.0	1.5
Liver cirrhosis	7.5	3.4	8.0	4.7
Lupus kidney	13.3	6.5	12.2	5.4
NCI-H292 none	21.8	11.0	14.9	15.5
NCI-H292 IL-4	42.0	36.1	43.5	44.4
NCI-H292 IL-9	41.8	48.3	32.8	28.1
NCI-H292 IL-13	20.9	17.2	30.4	21.6
NCI-H292 IFN gamma	14.4	12.9	14.6	10.2
HPAEC none	0.0	0.0	0.0	0.0
HPAEC TNF alpha + IL-1 beta	0.5	0.9	1.1	1.0
Lung fibroblast none	4.5	2.4	2.6	4.5
Lung fibroblast TNF alpha + IL-1 beta	0.3	0.2	0.0	0.0
Lung fibroblast IL-4	14.6	6.3	9.0	8.8
Lung fibroblast IL-9	3.9	1.0	8.6	2.9
Lung fibroblast IL-13	8.7	5.6	5.9	3.5
Lung fibroblast IFN gamma	14.9	3.8	6.7	4.5
Dermal fibroblast CCD1070 rest	0.0	0.0	0.0	0.0
Dermal fibroblast CCD1070 TNF alpha	0.0	0.0	0.0	0.0
Dermal fibroblast CCD1070 IL-1 beta	0.0	0.0	0.0	0.0

Dermal fibroblast IFN gamma	0.0	0.0	0.0	0.0
Dermal fibroblast IL-4	0.2	0.8	0.8	0.0
IBD Colitis 2	0.3	0.0	0.4	0.4
IBD Crohn's	9.9	4.1	2.7	3.6
Colon	41.2	20.9	27.5	20.0
Lung	61.6	35.1	34.6	35.4
Thymus	100.0	100.0	100.0	100.0
Kidney	21.3	13.9	14.9	14.0

Table ABR. Panel 5 Islet

Tissue Name	Rel. Exp.(%) Ag2505, Run 248045752	Tissue Name	Rel. Exp.(%) Ag2505, Run 248045752
97457_Patient-02go_adipose	32.3	94709_Donor 2 AM - A_adipose	0.0
97476_Patient-07sk_skeletal muscle	8.2	94710_Donor 2 AM - B_adipose	0.0
97477_Patient-07ut_uterus	31.4	94711_Donor 2 AM - C_adipose	0.0
97478_Patient-07pl_placenta	3.5	94712_Donor 2 AD - A_adipose	0.0
99167_Bayer Patient 1	9.5	94713_Donor 2 AD - B_adipose	0.0
97482_Patient-08ut_uterus	90.1	94714_Donor 2 AD - C_adipose	0.0
97483_Patient-08pl_placenta	7.4	94742_Donor 3 U - A_Mesenchymal Stem Cells	0.0
97486_Patient-09sk_skeletal muscle	1.4	94743_Donor 3 U - B_Mesenchymal Stem Cells	0.0
97487_Patient-09ut_uterus	78.5	94730_Donor 3 AM - A_adipose	0.0
97488_Patient-09pl_placenta	0.6	94731_Donor 3 AM - B_adipose	0.0
97492_Patient-10ut_uterus	66.0	94732_Donor 3 AM - C_adipose	0.2
97493_Patient-10pl_placenta	3.1	94733_Donor 3 AD - A_adipose	0.0
97495_Patient-11go_adipose	28.3	94734_Donor 3 AD - B_adipose	0.0
97496_Patient-11sk_skeletal muscle	5.8	94735_Donor 3 AD - C_adipose	0.0
97497_Patient-11ut_uterus	35.4	77138_Liver_HepG2untreated	21.2
97498_Patient-11pl_placenta	2.0	73556_Heart_Cardiac stromal cells (primary)	0.0
97500_Patient-12go_adipose	35.1	81735_Small Intestine	36.9

97501_Patient-12sk_skeletal muscle	9.9	72409_Kidney_Proximal Convoluted Tubule	4.6
97502_Patient-12ut uterus	100.0	82685_Small intestine_Duodenum	27.0
97503_Patient-12pl_placenta	4.1	90650_Adrenal_Adrenocortical adenoma	0.3
94721_Donor 2 U - A_Mesenchymal Stem Cells	0.0	72410_Kidney_HRCE	7.2
94722_Donor 2 U - B_Mesenchymal Stem Cells	0.0	72411_Kidney_HRE	10.7
94723_Donor 2 U - C_Mesenchymal Stem Cells	0.0	73139_Uterus_Uterine smooth muscle cells	0.0

Table ABS. general oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag2505, Run 267145080	Rel. Exp.(%) Ag5113, Run 260280405	Rel. Exp.(%) Ag5124, Run 259936347
Colon cancer 1	7.7	3.7	9.3
Colon NAT 1	2.5	1.3	3.6
Colon cancer 2	26.4	0.9	9.5
Colon NAT 2	8.4	2.8	0.0
Colon cancer 3	38.2	5.1	20.0
Colon NAT 3	16.7	14.0	18.6
Colon malignant cancer 4	69.7	1.9	20.3
Colon NAT 4	8.7	8.7	4.4
Lung cancer 1	7.2	2.4	5.3
Lung NAT 1	8.0	6.0	16.5
Lung cancer 2	26.1	8.4	67.8
Lung NAT 2	17.4	15.2	69.3
Squamous cell carcinoma 3	19.6	12.8	76.8
Lung NAT 3	8.3	2.6	0.0
Metastatic melanoma 1	14.5	11.5	36.1
Melanoma 2	1.1	0.0	0.0
Melanoma 3	2.5	0.8	1.4
Metastatic melanoma 4	13.5	7.3	100.0
Metastatic melanoma 5	12.6	11.0	99.3
Bladder cancer 1	1.0	2.0	0.0
Bladder NAT 1	0.0	0.0	0.0
Bladder cancer 2	2.4	3.5	6.0

Bladder NAT 2	0.3	0.5	0.0
Bladder NAT 3	0.7	0.0	0.0
Bladder NAT 4	1.5	0.0	12.5
Prostate adenocarcinoma 1	100.0	100.0	0.0
Prostate adenocarcinoma 2	12.9	5.6	2.8
Prostate adenocarcinoma 3	15.1	1.7	6.8
Prostate adenocarcinoma 4	15.9	2.1	18.2
Prostate NAT 5	14.7	2.3	0.0
Prostate adenocarcinoma 6	13.1	2.0	9.2
Prostate adenocarcinoma 7	16.2	8.1	35.6
Prostate adenocarcinoma 8	4.5	2.3	0.0
Prostate adenocarcinoma 9	33.4	21.0	69.3
Prostate NAT 10	7.7	0.9	0.0
Kidney cancer 1	9.6	1.1	18.0
Kidney NAT 1	33.4	4.0	27.4
Kidney cancer 2	19.8	6.0	65.1
Kidney NAT 2	64.6	10.2	31.6
Kidney cancer 3	7.1	1.6	2.2
Kidney NAT 3	29.7	1.7	12.7
Kidney cancer 4	8.5	5.2	12.8
Kidney NAT 4	7.9	2.1	26.6

AI_comprehensive panel_v1.0 Summary: Ag2505/Ag2831 Two experiments with different probes and primer sets are in excellent agreement, with highest expression of this gene seen in rheumatoid arthritis bone (CT=27-29). This gene shows ubiquitous expression, but expression of this gene is higher in bone, synovium, cartilage and synovial fluid from RA patients as compared to expression in samples from OA patients, normal and diseased lung. Expression of this gene is downregulated in Crohn's samples as compared to the corresponding control samples. This gene encode a putative novel adhesion molecule which is homologous to mouse POEM (preosteoblast epidermal growth factor-like repeat protein with meprin) or nephronectin. Murine nephronectin may function in multiple biological processes including development of the kidney (1) and bone (2) and contribute to liver and lung fibrosis (3). Therefore, therapeutic modulation of this gene may be useful in the treatment of autoimmune and inflammatory diseases such as rheumatoid and

osteoarthritis, Inflammatory bowel disease, COPD, asthma, psoriasis, liver and lung fibrosis.

References:

1. Miner JH. J Cell Biol 2001 Jul 23;154(2):257-9, PMID: 11470814.
- 5 2. Morimura N *et al.*, 2001, J. Biol. Chem. 2000 Nov 9;276(45):42172-42181, PMID: 11546798.
3. Levine *et al.*, 2000, Am J Pathol 2000 Jun;156(6):1927-35, PMID: 10854216.

CNS_neurodegeneration_v1.0

Summary: Ag2505/Ag2667/Ag2767/Ag2831/Ag7237 Six experiments with three different
 10 probe and primer sets are in excellent agreement. This panel confirms the expression of this gene at low levels in the brain in an independent group of individuals. This gene is found to be upregulated in the temporal cortex of Alzheimer's disease patients. This gene codes for a homolog of mouse POEM (Nephronectin short isoform), a cell adhesion molecule with EGF domains. Alpha secretase activity, which is generally believed to be a beneficial
 15 processing alternative to beta secretase, is increased by EGF in neuronal cells (1). This suggests the increased expression of this gene observed here is a compensatory action in the brain to counter the mechanisms of Alzheimer's Disease. Therefore, the protein encoded by this gene may be a potential therapeutic agent for the treatment of Alzheimer's disease and other neurodegenerative diseases.

20 EGF is also known to facilitate long term potentiation (LTP) in the hippocampus, a process thought to underlie learning and memory (2). Therefore, this gene may have utility in treating disorders of memory, such as neurodegenerative diseases and aging, when used alone or in combination with other growth factors such as bFGF.

In addition, EGF supports the growth and differentiation of dopaminergic neurons
 25 (3), which are selectively vulnerable to loss in Parkinson's disease. Therefore, this gene product may have utility in treating Parkinson's Disease.

Ag5113 Expression of this gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

References:

- 30 1. Slack BE, Breu J, Muchnicki L, Wurtman RJ, 1997, Biochem J 327 (Pt 1):245-9.

2. Abe K, Ishiyama J, Saito H, 1992, Brain Res 593(2):335-8.

3. Storch A, Paul G, Csete M, Boehm BO, Carvey PM, Kupsch A, Schwarz J, 2001, Exp Neurol 170(2):317-25.

General_screening_panel_v1.5 Summary: Ag5113/Ag5124 Highest expression
5 of this gene is detected in fetal lung (CT=29). Low but significant expression of this gene is also seen in tissues with metabolic function including adipose, pancreas, and gastrointestinal tract. See panel 1.3 for further discussion of this gene.

General_screening_panel_v1.6 Summary: Ag7237 Highest expression of this
gene is detected in fetal lung (CT=27). Expression of this gene is higher in fetal (CTs=27-
10 33) as compared to corresponding adult lung, kidney, liver and skeletal muscle tissues (CT=32-40). Therefore, expression of this gene may be used to distinguish between these fetal and adult tissues. In addition, the relative overexpression of this gene in fetal tissue suggests that the protein product may enhance growth or development of these tissues in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic
15 modulation of the protein encoded by this gene could be useful in treatment of lung, liver, kidney and muscle related diseases.

Moderate to low levels of expression of this gene is also seen in cancer cell lines derived from squamous cell carcinoma, brain, colon, renal, lung, breast, and ovarian cancers. Therefore, expression of this gene may be useful as diagnostic marker for detection
20 of these cancers. Furthermore, therapeutic modulation of this gene may be useful in the treatment of squamous cell carcinoma, brain, colon, renal, lung, breast, and ovarian cancers.

Moderate to low levels of expression of this gene is also seen in tissues with metabolic/endocrine functions and also in all the regions of brain. See panel 1.3D for further discussion of this gene.

Panel 1.3D Summary: Ag2505/Ag2667/Ag2767/Ag2831 Four experiments with
25 two different probes and primer sets are in good agreement. Highest expression of this gene is detected in the thyroid and fetal lung (CTs=29-31). Moderate to low levels of expression of this gene is also seen in other tissues with metabolic/endocrine functions, including skeletal muscle, fetal skeletal muscle, small intestine, stomach, pancreas, adipose and fetal
30 heart. Very low levels are also seen in heart and placenta. Nephronectin is the ligand for the $\alpha 8 \beta 1$ integrin as evidenced by two independent sets of published data (1,4). Integrins

are known to mediate development and organogenesis (5,6). Nephronectin can bind integrins including alpha5beta3, alpha5beta5, alpha5beta6 and alpha4beta7, but not alpha4beta1, alpha3beta1, alpha2beta1 or alpha1beta1. Nephronectin interacts with integrins via the RGD sequence, but RGD alone is not sufficient for binding, the MAM domain is
5 also required (2). MAM domains are thought to have an adhesive function. Thus, modulation of the expression or activity of this gene product by protein or antibody therapeutics may be an effective therapeutic for disorders involving alpha8beta1 integrin signaling including inflammatory diseases.

Obesity has also been linked as an inflammatory condition (7) and thus humanized
10 antibodies may also be therapeutically relevant in treating this condition and related complications such as type II diabetes.

Overall, this gene is expressed at a low to moderate level in the normal tissues on this panel. Furthermore, the brain, prostate, lung and colon cancer cell lines show a very low level of expression compared to the normal organs. This suggests that this molecule can
15 potentially be used as a therapeutic inhibitor for these cancers.

Moderate to low levels of expression is seen in all the regions of the central nervous system including substantia nigra, hippocampus, cortex, amygdala, thalamus and spinal cord. POEM is a ligand for alpha8beta1 integrin, which in turn promotes attachment, cell spreading, and neurite outgrowth on fibronectin (8). See CNS_neurodegeneration_v1.0 for
20 discussion of this gene in the central nervous system.

Reference:

4. Brandenberger R *et al.*, 2001, J Cell Biol 154(2):447-58, PMID: 11470831.
5. Schwartz *et al.*, 1995, Annu. Rev. Cell Dev. Biol. 11, 549-599, PMID: 8689569.
6. Clark and Brugge, 1995, Science 268, 233-239, PMID: 7716514.
- 25 7. Das UN, 2001, Nutrition 17(11-12):953-66, PMID: 11744348.
8. Muller *et al.*, 1995, Mol Biol Cell 6(4):433-48, PMID: 7626807

Panel 2.2 Summary: Ag2831 Highest expression of this gene is detected in kidney (CT=30.3). Expression of this gene is down regulated in kidney, lung and colon cancer as compared to the corresponding normal adjacent tissue. Conversely, increased expression of
30 this gene is seen in breast cancer samples. Therefore, expression of this gene may be used to

distinguish between cancer and normal kidney, lung, colon and breast. In addition, therapeutic modulation of this gene or its protein product in the form of protein therapeutic or through the use of antibodies may be useful in the treatment of kidney, lung, colon and breast cancer.

5 **Panel 2D Summary:** Ag2667/Ag2767/Ag2831 Three experiments with same probe and primer sets are in excellent agreement, with highest expression of this gene in metastatic breast cancer sample (CTs=26). Expression of this gene in this panel correlates with the expression pattern seen in panel 2.2. See panel 2.2 for further discussion of this gene.

10 **Panel 3D Summary:** Ag2831 Highest expression of this gene is detected in a small cell lung cancer NCI-H146 cell line (CT=29.7). Moderate to low levels of expression of this gene is also seen in cancer cell lines derived from epidermoid carcinoma, rhabdomyosarcoma, gastric, colon and small cell lung cancers. Therefore, expression of this gene may be used as diagnostic marker for detection of these cancers. Furthermore,
15 therapeutic modulation of this gene or its protein product through the use of antibodies may be useful in the treatment of these cancers.

Panel 4.1D Summary: Ag2831 Highest expression of this gene is detected in kidney (CT=31.3). In addition, moderate to low levels of expression of this gene is mainly seen in lung fibroblasts, and mucoepidermoid NCI-H292 cells. Expression of this gene is
20 upregulated in cytokine treated NCI-H292 cells, small airway epithelium and astrocytes. This expression pattern correlates with the expression observed in panel 4D. See panel 4D and AI panel for further discussion of this gene.

 Ag5113/Ag5124 Highest expression of this gene is seen in lung (CT=33). Low levels of expression of this gene is also seen in kidney and IL-4 treated lung fibroblasts.

25 **Panel 4D Summary:** Ag 2505 Highest expression of this transcript is found in the thymus and the lung(CTs=27-28). Consistent with this lung expression, this transcript is found in the pulmonary mucoepidermoid cell line H292 and is up-regulated upon treatment with the Th2 cytokines IL4 and IL9. This gene is also expressed at lower levels in lung fibroblasts treated with IL4. This transcript profile suggests that modulation of the
30 expression or activity of this gene product by protein or antibody therapeutics may be

beneficial for the treatment of inflammatory lung diseases such as asthma, emphysema and chronic obstructive pulmonary diseases.

Furthermore, therapeutics designed with the protein encoded for by this transcript could be important for maintaining or restoring normal function of thymus during inflammation.

Panel 5 Islet Summary: Ag2505 Highest expression of this gene is detected in uterus (CT=30). Moderate expression of this gene is also seen in adipose and skeletal muscle of gestational diabetic patients requiring and not requiring daily injections of insulin. This gene is also expressed in samples derived from pregnant and a nondiabetic, but overweight patient. In addition, this gene is also expressed in islet beta cells (those that are insulin producing) and small intestine. Therefore, therapeutic modulation of this gene may be useful in the treatment of metabolically related diseases including obesity, Type I and Type II diabetes.

general oncology screening panel_v_2.4 Summary: Ag2505 Highest expression of this gene is detected in prostate cancer (CT=27.7). Moderate to low levels of expression of this gene is seen in both normal and cancer samples derived from colon, lung, prostate and kidney. As Consistent with panels 2.2 and 2D, expression of this gene is downregulated in kidney cancer as compared to normal kidney. But higher expression of this gene is seen in colon cancer as compared to corresponding normal adjacent sample. Therefore, expression of this gene may be used to distinguish between cancer and normal kidney and colon tissue. See panel 1.3, 1.6, 2.2 for further discussion of this gene.

Ag5113/Ag5124 Highest expression of this gene is seen in metastatic melanoma and prostate cancer (CTs=31-33.7). Significant expression of this gene is seen in cancer samples derived from kidney, lung, and prostate cancers.

25 **AC. CG51264-01, CG51264-06 and CG51264-07: ST7-LIKE PROTEIN (17941787).**

Expression of gene CG51264-01, CG51264-06 and CG51264-07 was assessed using the primer-probe set Ag7547, described in Table ACA. Results of the RTQ-PCR runs are shown in Table ACB.

30 Table ACA. Probe Name Ag7547

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agcattgggatgtacttgtaagc-3'	23	1592	363
Probe	TET-5'- ctgtgtttcaaatgatcttctttcaaac a-3'-TAMRA	29	1630	364
Reverse	5'-ttctgcttccactcttgacaa-3'	21	1659	365

Table ACB. Panel 5 Islet

Tissue Name	Rel. Exp.(%) Ag7547, Run 308743747	Tissue Name	Rel. Exp.(%) Ag7547, Run 308743747
97457_Patient-02go_adipose	3.5	94709_Donor 2 AM - A_adipose	23.3
97476_Patient-07sk_skeletal muscle	0.0	94710_Donor 2 AM - B_adipose	23.0
97477_Patient-07ut_uterus	6.1	94711_Donor 2 AM - C_adipose	16.4
97478_Patient-07pl_placenta	1.0	94712_Donor 2 AD - A_adipose	43.5
99167_Bayer Patient I	4.8	94713_Donor 2 AD - B_adipose	66.0
97482_Patient-08ut_uterus	3.7	94714_Donor 2 AD - C_adipose	47.0
97483_Patient-08pl_placenta	1.4	94742_Donor 3 U - A_Mesenchymal Stem Cells	21.9
97486_Patient-09sk_skeletal muscle	7.3	94743_Donor 3 U - B_Mesenchymal Stem Cells	27.7
97487_Patient-09ut_uterus	4.6	94730_Donor 3 AM - A_adipose	41.2
97488_Patient-09pl_placenta	0.8	94731_Donor 3 AM - B_adipose	43.5
97492_Patient-10ut_uterus	8.0	94732_Donor 3 AM - C_adipose	47.0
97493_Patient-10pl_placenta	3.3	94733_Donor 3 AD - A_adipose	82.4
97495_Patient-11go_adipose	1.7	94734_Donor 3 AD - B_adipose	100.0
97496_Patient-11sk_skeletal muscle	5.9	94735_Donor 3 AD - C_adipose	31.9
97497_Patient-11ut_uterus	14.4	77138_Liver_HepG2untreated	4.5
97498_Patient-11pl_placenta	1.1	73556_Heart_Cardiac stromal cells (primary)	0.5
97500_Patient-12go_adipose	3.7	81735_Small Intestine	4.8
97501_Patient-12sk_skeletal muscle	14.2	72409_Kidney_Proximal Convoluted Tubule	15.5

97502_Patient-12ut_uterus	18.2	82685_Small intestine_Duodenum	3.1
97503_Patient-12pl_placenta	3.0	90650_Adrenal_Adrenocortical adenoma	0.8
94721_Donor 2 U - A_Mesenchymal Stem Cells	30.1	72410_Kidney_HRCE	49.0
94722_Donor 2 U - B_Mesenchymal Stem Cells	39.2	72411_Kidney_HRE	9.5
94723_Donor 2 U - C_Mesenchymal Stem Cells	51.1	73139_Uterus_Uterine smooth muscle cells	23.0

Panel 5 Islet Summary: Ag7547 Highest expression of this gene is detected in differentiated adipose tissue. Moderate levels of expression of this gene is mesenchymal stem cells, midway differentiated and differentiated adipose tissue. Low to moderate levels of expression of this gene is also detected in uterine smooth muscle, skeletal muscle from diabetic patient on insulin and kidney. Therefore, therapeutic modulation of this gene may be useful in the treatment of metabolic related diseases such as obesity, and diabetes.

AD. CG51264-03, and CG51264-04: (17941787-31) ST7-LIKE PROTEIN.

Expression of gene CG51264-03 and CG51264-04 was assessed using the primer-probe sets Ag2725 and Ag2727, described in Tables ADA and ADB.

Table ADA. Probe Name Ag2725

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctgcaactaccagaatcattgc-3'	22	1415	366
Probe	TET-5'- tggcaaacagaacccatctacttgggt -3'-TAMRA	26	1442	367
Reverse	5'-tgcaaggggatttaatgctact-3'	22	1469	368

Table ADB. Probe Name Ag2727

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctgcaactaccagaatcattgc-3'	22	1415	369
Probe	TET-5'- tggcaaacagaacccatctacttgggt -3'-TAMRA	26	1442	370

Reverse	5'-tgcaaggggatttaatgctact-3'	22	1469	371
---------	------------------------------	----	------	-----

AE. CG52423-01: PV1-LIKE PROTEIN (3544179_EXT).

- Expression of gene CG52423-01 was assessed using the primer-probe sets Ag1039, Ag1537, Ag760 and Ag4932, described in Tables AEA, AEB, AEC and AED. Results of the RTQ-PCR runs are shown in Tables AEE, AEF, AEG, AEH, AEI, AEJ, AEK, AEL, AEM and AEN.

Table AEA. Probe Name Ag1039

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-aaggagcaactgcaaaagg-3'	20	753	372
Probe	TET-5'-ctgcccctggacaaggacaagttt-3'-TAMRA	24	786	373
Reverse	5'-acagggttacgaagggtccatctc-3'	22	810	374

Table AEB. Probe Name Ag1537

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-aaggagctggaagagaagaaga-3'	22	1197	375
Probe	TET-5'-atcagaaactcagccctggacacctg-3'-TAMRA	26	1251	376
Reverse	5'-gctgcgacttggtcttgat-3'	19	1278	377

Table AEC. Probe Name Ag760

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-caccatgacaacgacacctata-3'	22	1924	378
Probe	TET-5'-atatggcaccaacatcacatgcacg-3'-TAMRA	25	1947	379
Reverse	5'-tgggtagaaaagtgtgtgtgaaa-3'	22	1979	380

Table AED. Probe Name Ag4932

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-aatgcagagatcaattcaagga-3'	22	535	381

Probe	TET-5'- aacaagagctgcatgaccttgcctt -3'-TAMRA	26	561	382
Reverse	5'-tcttcaccttctgattcagcat- 3'	22	588	383

Table AEE. Ardais Panel v.1.0

Tissue Name	Rel. Exp.(%) Ag1537, Run 267680189	Tissue Name	Rel. Exp.(%) Ag1537, Run 267680189
136799_Lung cancer(362)	23.8	136787_lung cancer(356)	8.1
136800_Lung NAT(363)	15.6	136788_lung NAT(357)	52.5
136813_Lung cancer(372)	45.4	136806_Lung cancer(36B)	35.6
136814_Lung NAT(373)	14.4	136807_Lung NAT(36C)	18.8
136815_Lung cancer(374)	39.2	136789_lung cancer(358)	65.1
136816_Lung NAT(375)	100.0	136802_Lung cancer(365)	49.3
136791_Lung cancer(35A)	22.5	136803_Lung cancer(368)	24.5
136795_Lung cancer(35E)	35.4	136804_Lung cancer(369)	38.2
136797_Lung cancer(360)	22.4	136811_Lung cancer(370)	14.9
136794_lung NAT(35D)	14.3	136810_Lung NAT(36F)	31.4
136818_Lung NAT(377)	33.0		

Table AEF. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag1537, Run 266937073	Rel. Exp.(%) Ag4932, Run 269217367	Tissue Name	Rel. Exp.(%) Ag1537, Run 266937073	Rel. Exp.(%) Ag4932, Run 269217367
AD 1 Hippo	13.1	9.5	Control (Path) 3 Temporal Ctx	4.2	0.0
AD 2 Hippo	14.2	22.1	Control (Path) 4 Temporal Ctx	1.5	5.7
AD 3 Hippo	0.0	3.4	AD 1 Occipital Ctx	5.9	6.8
AD 4 Hippo	3.5	1.9	AD 2 Occipital Ctx (Missing)	0.0	0.0
AD 5 Hippo	23.3	25.7	AD 3 Occipital Ctx	0.0	1.6
AD 6 Hippo	16.6	29.5	AD 4 Occipital Ctx	5.4	4.2
Control 2 Hippo	43.8	28.1	AD 5 Occipital Ctx	25.2	18.8
Control 4 Hippo	100.0	56.6	AD 6 Occipital Ctx	4.3	6.9

Control (Path) 3 Hippo	49.3	100.0	Control 1 Occipital Ctx	0.0	0.0
AD 1 Temporal Ctx	11.5	8.3	Control 2 Occipital Ctx	19.3	14.0
AD 2 Temporal Ctx	28.5	25.3	Control 3 Occipital Ctx	23.5	8.2
AD 3 Temporal Ctx	1.7	0.9	Control 4 Occipital Ctx	3.3	4.1
AD 4 Temporal Ctx	3.8	11.7	Control (Path) 1 Occipital Ctx	15.4	13.5
AD 5 Inf Temporal Ctx	31.0	36.3	Control (Path) 2 Occipital Ctx	7.9	1.1
AD 5 Sup Temporal Ctx	67.8	96.6	Control (Path) 3 Occipital Ctx	0.0	1.0
AD 6 Inf Temporal Ctx	23.7	38.2	Control (Path) 4 Occipital Ctx	0.0	9.0
AD 6 Sup Temporal Ctx	13.3	22.4	Control 1 Parietal Ctx	3.4	0.8
Control 1 Temporal Ctx	0.0	6.3	Control 2 Parietal Ctx	23.7	22.2
Control 2 Temporal Ctx	34.2	28.7	Control 3 Parietal Ctx	4.0	0.0
Control 3 Temporal Ctx	12.9	13.4	Control (Path) 1 Parietal Ctx	28.3	14.0
Control 3 Temporal Ctx	13.0	6.8	Control (Path) 2 Parietal Ctx	5.0	10.0
Control (Path) 1 Temporal Ctx	43.5	26.1	Control (Path) 3 Parietal Ctx	0.0	1.2
Control (Path) 2 Temporal Ctx	12.2	10.0	Control (Path) 4 Parietal Ctx	16.3	12.2

Table AEG. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag4932, Run 228843451	Tissue Name	Rel. Exp.(%) Ag4932, Run 228843451
Adipose	0.0	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	86.5
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	4.5
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca. * (SW480 met) SW620	0.0
Testis Pool	10.7	Colon ca. HT29	0.0
Prostate ca. * (bone met) PC-3	0.0	Colon ca. HCT-116	0.0

Prostate Pool	14.7	Colon ca. CaCo-2	0.0
Placenta	42.6	Colon cancer tissue	51.1
Uterus Pool	53.6	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	89.5
Ovarian ca. OVCAR-5	0.2	Small Intestine Pool	11.7
Ovarian ca. IGROV-1	0.0	Stomach Pool	37.9
Ovarian ca. OVCAR-8	0.1	Bone Marrow Pool	46.0
Ovary	11.1	Fetal Heart	15.1
Breast ca. MCF-7	0.0	Heart Pool	22.2
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	66.9
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	23.2
Breast ca. T47D	0.0	Skeletal Muscle Pool	32.5
Breast ca. MDA-N	0.0	Spleen Pool	100.0
Breast Pool	62.4	Thymus Pool	30.6
Trachea	47.3	CNS cancer (glio/astro) U87-MG	0.0
Lung	4.3	CNS cancer (glio/astro) U-118-MG	0.1
Fetal Lung	17.2	CNS cancer (neuro;met) SK-N-AS	0.1
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF- 539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB- 75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB- 19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.7
Lung ca. A549	0.0	Brain (Amygdala) Pool	1.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	2.0
Lung ca. NCI-H23	0.0	Brain (fetal)	2.6
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	2.1
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	1.3
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	1.5
Liver	1.2	Brain (Thalamus) Pool	2.5
Fetal Liver	44.1	Brain (whole)	3.8
Liver ca. HepG2	0.0	Spinal Cord Pool	1.9
Kidney Pool	55.1	Adrenal Gland	55.1
Fetal Kidney	73.2	Pituitary gland Pool	10.3
Renal ca. 786-0	0.0	Salivary Gland	20.7
Renal ca. A498	0.1	Thyroid (female)	70.7
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0

Renal ca. UO-31	0.0	Pancreas Pool	53.6
-----------------	-----	---------------	------

Table AEH. Oncology_cell_line_screening_panel_v3.2

Tissue Name	Rel. Exp.(%) Ag1537, Run 267177 741	Tissue Name	Rel. Exp.(%) Ag1537, Run 267177 41
94905_Daoy_Medulloblastoma/Cerebellum_sscDNA	0.0	94954_Ca Ski_Cervical epidermoid carcinoma (metastasis)_sscDNA	0.2
94906_TE671_Medulloblastom/Cerebellum_sscDNA	0.0	94955_ES-2_Ovarian clear cell carcinoma_sscDNA	0.0
94907_D283_Med_Medulloblastoma/Cerebellum_sscDNA	0.0	94957_Ramos/6h stim_Stimulated with PMA/ionomycin 6h_sscDNA	0.0
94908_PFSK-1_Primitive Neuroectodermal/Cerebellum_sscDNA	0.5	94958_Ramos/14h stim_Stimulated with PMA/ionomycin 14h_sscDNA	0.0
94909_XF-498_CNS_sscDNA	0.0	94962_MEG-01_Chronic myelogenous leukemia (megakaryoblast)_sscDNA	0.7
94910_SNB-78_CNS/glioma_sscDNA	0.0	94963_Raji_Burkitt's lymphoma_sscDNA	0.0
94911_SF-268_CNS/glioblastoma_sscDNA	0.0	94964_Daudi_Burkitt's lymphoma_sscDNA	0.8
94912_T98G_Glioblastoma_sscDNA	0.0	94965_U266_B-cell plasmacytoma/myeloma_sscDNA	1.3
96776_SK-N-SH_Neuroblastoma (metastasis)_sscDNA	0.0	94968_CA46_Burkitt's lymphoma_sscDNA	0.2
94913_SF-295_CNS/glioblastoma_sscDNA	0.0	94970_RL_non-Hodgkin's B-cell lymphoma_sscDNA	0.0
132565_NT2 pool_sscDNA	0.1	94972_JM1_pre-B-cell lymphoma/leukemia_sscDNA	0.0
94914_Cerebellum_sscDNA	0.2	94973_Jurkat_T cell leukemia_sscDNA	0.0
96777_Cerebellum_sscDNA	0.3	94974_TF-1_Erythroleukemia_sscDNA	100.0
94916_NCI-H292_Mucoepidermoid lung carcinoma_sscDNA	0.0	94975_HUT 78_T-cell lymphoma_sscDNA	0.0
94917_DMS-114_Small cell lung cancer_sscDNA	0.0	94977_U937_Histiocytic lymphoma_sscDNA	0.0
94918_DMS-79_Small cell lung cancer/neuroendocrine_sscDNA	0.0	94980_KU-812_Myelogenous leukemia_sscDNA	28.9
94919_NCI-H146_Small cell lung cancer/neuroendocrine_sscDNA	0.0	94981_769-P_Clear cell renal carcinoma_sscDNA	0.1

94920_NCI-H526_Small cell lung cancer/neuroendocrine_sscDNA	0.0	94983_Caki-2_Clear cell renal carcinoma_sscDNA	0.0
94921_NCI-N417_Small cell lung cancer/neuroendocrine_sscDNA	0.0	94984_SW 839_Clear cell renal carcinoma_sscDNA	0.0
94923_NCI-H82_Small cell lung cancer/neuroendocrine_sscDNA	0.0	94986_G401_Wilms' tumor_sscDNA	0.0
94924_NCI-H157_Squamous cell lung cancer (metastasis)_sscDNA	0.0	126768_293 cells_sscDNA	0.0
94925_NCI-H1155_Large cell lung cancer/neuroendocrine_sscDNA	0.0	94987_Hs766T_Pancreatic carcinoma (LN metastasis)_sscDNA	0.6
94926_NCI-H1299_Large cell lung cancer/neuroendocrine_sscDNA	0.0	94988_CAPAN-1_Pancreatic adenocarcinoma (liver metastasis)_sscDNA	0.0
94927_NCI-H727_Lung carcinoid_sscDNA	0.0	94989_SU86.86_Pancreatic carcinoma (liver metastasis)_sscDNA	0.0
94928_NCI-UMC-11_Lung carcinoid_sscDNA	0.0	94990_BxPC-3_Pancreatic adenocarcinoma_sscDNA	0.0
94929_LX-1_Small cell lung cancer_sscDNA	0.0	94991_HPAC_Pancreatic adenocarcinoma_sscDNA	0.0
94930_Colo-205_Colon cancer_sscDNA	0.0	94992_MIA PaCa-2_Pancreatic carcinoma_sscDNA	0.0
94931_KM12_Colon cancer_sscDNA	0.0	94993_CFPAC-1_Pancreatic ductal adenocarcinoma_sscDNA	0.1
94932_KM20L2_Colon cancer_sscDNA	0.0	94994_PANC-1_Pancreatic epithelioid ductal carcinoma_sscDNA	0.0
94933_NCI-H716_Colon cancer_sscDNA	0.0	94996_T24_Bladder carcinoma (transitional cell)_sscDNA	0.1
94935_SW-48_Colon adenocarcinoma_sscDNA	0.0	94997_5637_Bladder carcinoma_sscDNA	0.0
94936_SW1116_Colon adenocarcinoma_sscDNA	0.0	94998_HT-1197_Bladder carcinoma_sscDNA	0.0
94937_LS 174T_Colon adenocarcinoma_sscDNA	0.0	94999_UM-UC-3_Bladder carcinoma (transitional cell)_sscDNA	0.0
94938_SW-948_Colon adenocarcinoma_sscDNA	0.0	95000_A204_Rhabdomyosarcoma_sscDNA	0.0
94939_SW-480_Colon adenocarcinoma_sscDNA	0.0	95001_HT-1080_Fibrosarcoma_sscDNA	0.0
94940_NCI-SNU-5_Gastric carcinoma_sscDNA	0.0	95002_MG-63_Osteosarcoma (bone)_sscDNA	0.0
112197_KATO III_Stomach_sscDNA	0.0	95003_SK-LMS-1_Leiomyosarcoma (vulva)_sscDNA	0.2
94943_NCI-SNU-16_Gastric carcinoma_sscDNA	0.0	95004_SJRH30_Rhabdomyosarcoma (met to bone marrow)_sscDNA	0.0
94944_NCI-SNU-1_Gastric carcinoma_sscDNA	0.0	95005_A431_Epidermoid carcinoma_sscDNA	0.0

94946_RF-1_Gastric adenocarcinoma_sscDNA	0.0	95007_WM266-4_Melanoma_sscDNA	0.0
94947_RF-48_Gastric adenocarcinoma_sscDNA	0.1	112195_DU 145_Prostate_sscDNA	0.0
96778_MKN-45_Gastric carcinoma_sscDNA	0.0	95012_MDA-MB-468_Breast adenocarcinoma_sscDNA	0.0
94949_NCI-N87_Gastric carcinoma_sscDNA	0.0	112196_SSC-4_Tongue_sscDNA	0.0
94951_OVCAR-5_Ovarian carcinoma_sscDNA	0.0	112194_SSC-9_Tongue_sscDNA	0.0
94952_RL95-2_Uterine carcinoma_sscDNA	0.0	112191_SSC-15_Tongue_sscDNA	0.0
94953_HelaS3_Cervical adenocarcinoma_sscDNA	0.0	95017_CAL 27_Squamous cell carcinoma of tongue_sscDNA	0.0

Table AEI. Panel 1.2

Tissue Name	Rel. Exp.(%) Ag1537, Run 142331743	Rel. Exp.(%) Ag760, Run 114246835	Tissue Name	Rel. Exp.(%) Ag1537, Run 142331743	Rel. Exp.(%) Ag760, Run 114246835
Endothelial cells	2.5	1.3	Renal ca. 786-0	0.0	0.0
Heart (Fetal)	17.6	2.3	Renal ca. A498	0.1	0.1
Pancreas	35.4	74.2	Renal ca. RXF 393	0.0	0.0
Pancreatic ca. CAPAN 2	0.0	0.0	Renal ca. ACHN	0.0	0.0
Adrenal Gland	37.4	19.1	Renal ca. UO-31	0.0	0.0
Thyroid	14.9	100.0	Renal ca. TK-10	0.0	0.0
Salivary gland	34.6	15.8	Liver	2.1	1.6
Pituitary gland	2.1	27.4	Liver (fetal)	4.4	4.0
Brain (fetal)	0.1	0.7	Liver ca. (hepatoblast) HepG2	0.0	0.1
Brain (whole)	0.2	0.5	Lung	1.0	4.1
Brain (amygdala)	0.3	0.3	Lung (fetal)	0.3	2.1
Brain (cerebellum)	0.1	0.1	Lung ca. (small cell) LX-1	0.0	0.0
Brain (hippocampus)	0.8	0.7	Lung ca. (small cell) NCI-H69	0.0	0.0
Brain (thalamus)	0.6	0.4	Lung ca. (s.cell var.) SHP-77	0.0	0.0
Cerebral Cortex	0.8	0.3	Lung ca. (large cell) NCI-H460	0.0	0.0
Spinal cord	0.1	0.6	Lung ca. (non-sm. cell) A549	0.0	0.0
glio/astro U87-MG	0.0	0.0	Lung ca. (non-s.cell) NCI-H23	0.0	0.0

glio/astro U-118-MG	0.0	0.0	Lung ca. (non-s.cell) HOP-62	0.0	0.0
astrocytoma SW1783	0.0	0.0	Lung ca. (non-s.cl) NCI-H522	0.0	0.0
neuro*; met SK-N- AS	0.0	0.0	Lung ca. (squam.) SW 900	0.0	0.0
astrocytoma SF-539	0.0	0.0	Lung ca. (squam.) NCI-H596	0.0	0.0
astrocytoma SNB-75	0.0	0.0	Mammary gland	14.8	19.3
glioma SNB-19	0.0	0.0	Breast ca.* (pl.ef) MCF-7	0.0	0.0
glioma U251	0.1	0.2	Breast ca.* (pl.ef) MDA-MB-231	0.0	0.0
glioma SF-295	0.1	0.1	Breast ca.* (pl. ef) T47D	0.1	0.0
Heart	50.3	17.0	Breast ca. BT-549	0.0	0.0
Skeletal Muscle	18.2	16.0	Breast ca. MDA-N	2.2	1.2
Bone marrow	2.7	1.4	Ovary	3.0	0.8
Thymus	0.9	2.8	Ovarian ca. OVCAR-3	0.0	0.0
Spleen	29.1	30.8	Ovarian ca. OVCAR-4	0.0	0.0
Lymph node	2.7	14.4	Ovarian ca. OVCAR-5	0.1	0.1
Colorectal Tissue	2.3	1.1	Ovarian ca. OVCAR-8	0.2	0.1
Stomach	11.5	33.2	Ovarian ca. IGROV- I	0.0	0.0
Small intestine	52.5	41.5	Ovarian ca. (ascites) SK-OV-3	0.0	0.0
Colon ca. SW480	0.0	0.0	Uterus	9.2	12.8
Colon ca.* SW620 (SW480 met)	0.0	0.0	Placenta	3.1	7.3
Colon ca. HT29	0.0	0.0	Prostate	19.5	12.3
Colon ca. HCT-116	0.0	0.0	Prostate ca.* (bone met) PC-3	0.0	0.0
Colon ca. CaCo-2	0.0	0.0	Testis	0.2	1.4
Colon ca. Tissue (ODO3866)	1.7	1.4	Melanoma Hs688(A).T	0.0	0.0
Colon ca. HCC-2998	0.0	0.0	Melanoma* (met) Hs688(B).T	0.0	0.0
Gastric ca.* (liver met) NCI-N87	0.9	0.7	Melanoma UACC- 62	0.0	0.0
Bladder	52.5	13.1	Melanoma M14	0.0	0.0
Trachea	2.1	9.6	Melanoma LOX IMVI	0.0	0.0

Kidney	100.0	22.4	Melanoma* (met) SK-MEL-5	0.0	0.0
Kidney (fetal)	23.8	31.9			

Table AEJ. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag760, Run 165678100	Tissue Name	Rel. Exp.(%) Ag760, Run 165678100
Liver adenocarcinoma	0.0	Kidney (fetal)	33.4
Pancreas	43.8	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	0.2
Adrenal gland	21.5	Renal ca. RXF 393	0.0
Thyroid	79.6	Renal ca. ACHN	0.0
Salivary gland	13.9	Renal ca. UO-31	0.0
Pituitary gland	13.4	Renal ca. TK-10	0.0
Brain (fetal)	0.7	Liver	1.9
Brain (whole)	0.9	Liver (fetal)	12.4
Brain (amygdala)	1.6	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	0.4	Lung	15.3
Brain (hippocampus)	1.8	Lung (fetal)	6.1
Brain (substantia nigra)	2.3	Lung ca. (small cell) LX-1	0.0
Brain (thalamus)	2.7	Lung ca. (small cell) NCI- H69	0.0
Cerebral Cortex	0.7	Lung ca. (s.cell var.) SHP- 77	0.0
Spinal cord	1.7	Lung ca. (large cell) NCI- H460	0.4
glio/astro U87-MG	0.0	Lung ca. (non-sm. cell) A549	0.0
glio/astro U-118-MG	0.1	Lung ca. (non-s.cell) NCI- H23	0.0
astrocytoma SW1783	0.0	Lung ca. (non-s.cell) HOP-62	0.0
neuro*; met SK-N-AS	0.0	Lung ca. (non-s.cl) NCI- H522	0.0
astrocytoma SF-539	0.1	Lung ca. (squam.) SW 900	0.0
astrocytoma SNB-75	0.0	Lung ca. (squam.) NCI- H596	0.0
glioma SNB-19	0.0	Mammary gland	26.8
glioma U251	0.7	Breast ca.* (pl.ef) MCF-7	0.0
glioma SF-295	0.0	Breast ca.* (pl.ef) MDA- MB-231	0.0

Heart (fetal)	6.9	Breast ca.* (pl.ef) T47D	0.0
Heart	11.0	Breast ca. BT-549	0.0
Skeletal muscle (fetal)	19.5	Breast ca. MDA-N	0.2
Skeletal muscle	9.9	Ovary	1.8
Bone marrow	7.9	Ovarian ca. OVCAR-3	0.1
Thymus	6.9	Ovarian ca. OVCAR-4	0.0
Spleen	90.8	Ovarian ca. OVCAR-5	0.0
Lymph node	73.7	Ovarian ca. OVCAR-8	0.0
Colorectal	7.9	Ovarian ca. IGROV-1	0.0
Stomach	65.5	Ovarian ca.* (ascites) SK-OV-3	0.1
Small intestine	100.0	Uterus	87.7
Colon ca. SW480	0.0	Placenta	6.4
Colon ca.* SW620(SW480 met)	0.0	Prostate	11.3
Colon ca. HT29	0.0	Prostate ca.* (bone met)PC-3	0.0
Colon ca. HCT-116	0.0	Testis	2.1
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
Colon ca. tissue(ODO3866)	24.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	0.0	Melanoma UACC-62	0.0
Gastric ca.* (liver met) NCI-N87	1.7	Melanoma M14	0.0
Bladder	17.1	Melanoma LOX IMVI	0.0
Trachea	27.0	Melanoma* (met) SK-MEL-5	0.0
Kidney	18.2	Adipose	26.6

Table AEK. Panel 2D

Tissue Name	Rel. Exp.(%) Ag1537, Run 145017308	Tissue Name	Rel. Exp.(%) Ag1537, Run 145017308
Normal Colon	12.3	Kidney Margin 8120608	23.5
CC Well to Mod Diff (ODO3866)	10.7	Kidney Cancer 8120613	21.5
CC Margin (ODO3866)	12.2	Kidney Margin 8120614	12.3
CC Gr.2 rectosigmoid (ODO3868)	3.2	Kidney Cancer 9010320	34.4
CC Margin (ODO3868)	0.8	Kidney Margin 9010321	27.7
CC Mod Diff (ODO3920)	3.4	Normal Uterus	9.3
CC Margin (ODO3920)	2.2	Uterus Cancer 064011	6.4

CC Gr.2 ascend colon (ODO3921)	13.4	Normal Thyroid	84.1
CC Margin (ODO3921)	5.8	Thyroid Cancer 064010	20.6
CC from Partial Hepatectomy (ODO4309) Mets	9.6	Thyroid Cancer A302152	15.2
Liver Margin (ODO4309)	0.6	Thyroid Margin A302153	21.3
Colon mets to lung (OD04451-01)	5.5	Normal Breast	22.1
Lung Margin (OD04451-02)	0.8	Breast Cancer (OD04566)	8.4
Normal Prostate 6546-1	14.1	Breast Cancer (OD04590-01)	21.0
Prostate Cancer (OD04410)	8.8	Breast Cancer Mets (OD04590-03)	27.7
Prostate Margin (OD04410)	6.9	Breast Cancer Metastasis (OD04655-05)	9.1
Prostate Cancer (OD04720-01)	3.1	Breast Cancer 064006	10.1
Prostate Margin (OD04720-02)	10.3	Breast Cancer 1024	7.1
Normal Lung 061010	11.8	Breast Cancer 9100266	10.4
Lung Met to Muscle (ODO4286)	6.4	Breast Margin 9100265	7.4
Muscle Margin (ODO4286)	9.9	Breast Cancer A209073	27.4
Lung Malignant Cancer (OD03126)	19.3	Breast Margin A209073	8.7
Lung Margin (OD03126)	3.3	Normal Liver	1.1
Lung Cancer (OD04404)	5.2	Liver Cancer 064003	6.5
Lung Margin (OD04404)	25.3	Liver Cancer 1025	0.7
Lung Cancer (OD04565)	3.4	Liver Cancer 1026	8.1
Lung Margin (OD04565)	3.1	Liver Cancer 6004-T	1.9
Lung Cancer (OD04237-01)	11.0	Liver Tissue 6004-N	3.6
Lung Margin (OD04237-02)	18.2	Liver Cancer 6005-T	9.3
Ocular Mel Met to Liver (ODO4310)	0.7	Liver Tissue 6005-N	0.6

Liver Margin (ODO4310)	1.7	Normal Bladder	14.1
Melanoma Mets to Lung (OD04321)	3.9	Bladder Cancer 1023	4.5
Lung Margin (OD04321)	3.7	Bladder Cancer A302173	3.6
Normal Kidney	40.6	Bladder Cancer (OD04718-01)	7.4
Kidney Ca, Nuclear grade 2 (OD04338)	5.7	Bladder Normal Adjacent (OD04718-03)	15.2
Kidney Margin (OD04338)	11.1	Normal Ovary	1.4
Kidney Ca Nuclear grade 1/2 (OD04339)	2.5	Ovarian Cancer 064008	6.5
Kidney Margin (OD04339)	17.6	Ovarian Cancer (OD04768-07)	1.6
Kidney Ca, Clear cell type (OD04340)	100.0	Ovary Margin (OD04768-08)	9.2
Kidney Margin (OD04340)	22.7	Normal Stomach	13.5
Kidney Ca, Nuclear grade 3 (OD04348)	55.1	Gastric Cancer 9060358	2.8
Kidney Margin (OD04348)	19.9	Stomach Margin 9060359	12.6
Kidney Cancer (OD04622-01)	25.0	Gastric Cancer 9060395	20.6
Kidney Margin (OD04622-03)	7.4	Stomach Margin 9060394	7.5
Kidney Cancer (OD04450-01)	1.3	Gastric Cancer 9060397	10.0
Kidney Margin (OD04450-03)	9.2	Stomach Margin 9060396	3.2
Kidney Cancer 8120607	9.2	Gastric Cancer 064005	6.7

Table AEL. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4932, Run 223597251	Tissue Name	Rel. Exp.(%) Ag4932, Run 223597251
Secondary Th1 act	0.1	HUVEC IL-1beta	5.6
Secondary Th2 act	0.4	HUVEC IFN gamma	40.6
Secondary Tr1 act	0.1	HUVEC TNF alpha + IFN gamma	4.6

Secondary Th1 rest	0.1	HUVEC TNF alpha + IL4	5.0
Secondary Th2 rest	0.0	HUVEC IL-11	8.7
Secondary Tr1 rest	0.0	Lung Microvascular EC none	66.4
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	30.4
Primary Th2 act	0.0	Microvascular Dermal EC none	43.5
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	17.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.3
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	1.2	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.2	Coronary artery SMC TNFalpha + IL-1beta	1.1
CD8 lymphocyte act	0.1	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.4	Astrocytes TNFalpha + IL-1beta	0.2
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	27.0
CD4 lymphocyte none	0.3	KU-812 (Basophil) PMA/ionomycin	28.3
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.1	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.1	Liver cirrhosis	20.6
LAK cells IL-2+IL-12	0.2	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.5	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.2	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.2	NCI-H292 IL-13	0.1
NK Cells IL-2 rest	0.2	NCI-H292 IFN gamma	0.0

Two Way MLR 3 day	2.7	HPAEC none	1.8
Two Way MLR 5 day	1.3	HPAEC TNF alpha + IL-1 beta	1.5
Two Way MLR 7 day	0.1	Lung fibroblast none	0.4
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.6
PBMC PWM	0.0	Lung fibroblast IL-4	0.2
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.4
Ramos (B cell) ionomycin	0.1	Lung fibroblast IFN gamma	0.2
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.5	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.1	Dermal fibroblast IFN gamma	1.1
Dendritic cells none	0.1	Dermal fibroblast IL-4	0.4
Dendritic cells LPS	1.7	Dermal Fibroblasts rest	0.7
Dendritic cells anti-CD40	0.9	Neutrophils TNFa+LPS	0.4
Monocytes rest	0.6	Neutrophils rest	0.3
Monocytes LPS	0.1	Colon	19.3
Macrophages rest	0.0	Lung	100.0
Macrophages LPS	0.1	Thymus	48.0
HUVEC none	1.9	Kidney	68.8
HUVEC starved	8.4		

Table AEM. Panel 4D

Tissue Name	Rel. Exp.(%) Ag760, Run 145803954	Tissue Name	Rel. Exp.(%) Ag760, Run 145803954
Secondary Th1 act	0.0	HUVEC IL-1beta	3.4
Secondary Th2 act	0.1	HUVEC IFN gamma	36.6

Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	4.0
Secondary Th1 rest	0.1	HUVEC TNF alpha + IL4	3.4
Secondary Th2 rest	0.0	HUVEC IL-11	5.5
Secondary Tr1 rest	0.0	Lung Microvascular EC none	47.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1 beta	22.8
Primary Th2 act	0.0	Microvascular Dermal EC none	40.1
Primary Tr1 act	0.1	Microvascular Dermal EC TNFalpha + IL-1 beta	17.9
Primary Th1 rest	0.1	Bronchial epithelium TNFalpha + IL1 beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1 beta	0.0
CD45RA CD4 lymphocyte act	0.6	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.2	Coronary artery SMC TNFalpha + IL-1 beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1 beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	24.3
CD4 lymphocyte none	0.3	KU-812 (Basophil) PMA/ionomycin	29.7
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.1	CCD1106 (Keratinocytes) TNFalpha + IL-1 beta	0.0
LAK cells IL-2	0.1	Liver cirrhosis	19.5
LAK cells IL-2+IL-12	0.0	Lupus kidney	34.4
LAK cells IL-2+IFN gamma	1.0	NCI-H292 none	0.0
LAK cells IL-2+ IL-18	0.7	NCI-H292 IL-4	0.0

LAK cells PMA/ionomycin	0.0	NCI-H292 IL-9	0.0
NK Cells IL-2 rest	0.4	NCI-H292 IL-13	0.0
Two Way MLR 3 day	3.5	NCI-H292 IFN gamma	0.0
Two Way MLR 5 day	1.3	HPAEC none	0.9
Two Way MLR 7 day	0.1	HPAEC TNF alpha + IL-1 beta	0.7
PBMC rest	0.1	Lung fibroblast none	0.0
PBMC PWM	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PHA-L	0.1	Lung fibroblast IL-4	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) ionomycin	0.1	Lung fibroblast IL-13	0.0
B lymphocytes PWM	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes CD40L and IL-4	0.3	Dermal fibroblast CCD1070 rest	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.1
Dendritic cells none	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells LPS	2.3	Dermal fibroblast IL-4	0.1
Dendritic cells anti-CD40	0.0	IBD Colitis 2	1.5
Monocytes rest	0.8	IBD Crohn's	9.0
Monocytes LPS	0.0	Colon	40.3
Macrophages rest	0.0	Lung	100.0
Macrophages LPS	0.6	Thymus	95.3
HUVEC none	3.8	Kidney	59.9
HUVEC starved	16.8		

Table AEN. general oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag1537, Run 266930996	Rel. Exp.(%) Ag760, Run 262228031	Tissue Name	Rel. Exp.(%) Ag1537, Run 266930996	Rel. Exp.(%) Ag760, Run 262228031
-------------	--	--	-------------	--	---

Colon cancer 1	11.4	4.8	Bladder cancer NAT 2	0.6	0.4
Colon cancer NAT 1	9.4	1.7	Bladder cancer NAT 3	0.8	0.2
Colon cancer 2	4.1	3.9	Bladder cancer NAT 4	4.2	2.3
Colon cancer NAT 2	4.1	1.8	Prostate adenocarcinoma 1	2.1	2.9
Colon cancer 3	10.4	4.8	Prostate adenocarcinoma 2	1.2	0.5
Colon cancer NAT 3	7.6	1.2	Prostate adenocarcinoma 3	2.1	1.0
Colon malignant cancer 4	9.7	4.0	Prostate adenocarcinoma 4	6.0	3.3
Colon normal adjacent tissue 4	3.4	2.3	Prostate cancer NAT 5	3.0	1.1
Lung cancer 1	4.8	3.6	Prostate adenocarcinoma 6	1.3	0.5
Lung NAT 1	0.7	0.5	Prostate adenocarcinoma 7	1.2	0.7
Lung cancer 2	5.6	3.9	Prostate adenocarcinoma 8	1.2	0.4
Lung NAT 2	0.1	0.1	Prostate adenocarcinoma 9	4.8	2.7
Squamous cell carcinoma 3	5.4	2.4	Prostate cancer NAT 10	0.6	0.5
Lung NAT 3	0.7	0.3	Kidney cancer 1	90.1	100.0
metastatic melanoma 1	1.5	1.1	Kidney NAT 1	5.0	3.5
Melanoma 2	4.0	2.6	Kidney cancer 2	60.3	55.1
Melanoma 3	3.8	1.2	Kidney NAT 2	9.8	6.0
metastatic melanoma 4	2.2	0.9	Kidney cancer 3	30.8	39.5
metastatic melanoma 5	4.6	1.5	Kidney NAT 3	5.4	1.2
Bladder cancer 1	1.0	0.6	Kidney cancer 4	100.0	29.9
Bladder cancer NAT 1	0.0	0.0	Kidney NAT 4	6.7	1.6

Bladder cancer 2	4.4	2.1			
------------------	-----	-----	--	--	--

Ardais Panel v.1.0 Summary: Ag1537 Highest expression of this gene is detected in normal lung sample (CT=26.7). In addition, high to moderate levels of expression is seen in both cancer and normal lung samples. Therefore, therapeutic modulation of the PV1 protein (PLVAP) encoded by this gene may be useful in the treatment of certain subtypes of lung cancer.

CNS_neurodegeneration_v1.0 Summary: Ag1537/Ag4932 Two experiments with different probe and primer sets are in good agreement. This panel confirms the expression of this gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. See Panel 1.5 for a discussion of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.5 Summary: Ag4932 Highest expression of this gene is detected in spleen (CT=26). In addition, high expression of this gene is also detected in tissues with metabolic/endocrine functions including pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. The PV-1-like protein is a plasma membrane protein with an extracellular domain. The extracellular domain of this protein makes it a potential antibody target for the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

Moderate levels of expression of this gene is also seen in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, therapeutic modulation of this gene product may be useful in the treatment of central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

In addition, this gene also shows high expression in colon cancer tissue, with moderate levels of expression in a gastric NCI-N87 cell line. Therefore, therapeutic modulation of this gene may be useful in the treatment of colon and gastric cancers.

HASS Panel v1.0 Summary: Ag1537 Expression of this gene is low/undetectable (CTs > 34.9) across all of the samples on this panel (data not shown).

Oncology_cell_line_screening_panel_v3.2 Summary: Ag1537 Highest expression of this gene is detected in TF-1 erythroleukemia cells (CT=28.6). Moderate levels of expression of this gene is restricted to erythroleukemia and myelogenous leukemia. Therefore, expression of this gene may be used to distinguish these leukemia samples from
5 other samples in the panel and also, as marker to detect the presence of these leukemia. In addition, therapeutic modulation of this gene or its protein product may be useful in the treatment of erythroleukemia and myelogenous leukemia.

Panel 1.2 Summary: Ag760/Ag1537 Results from two experiments using different probe/primer sets are in reasonable agreement with highest expression of this gene in
10 thyroid and kidney (CTs=20-21.6). Expression of this gene seems to be restricted to normal tissue and it is low or undetectable in cancer cell lines. Thus, expression of this gene could be used to distinguish between normal tissues and cultured cancer cell lines.

In addition, expression of this gene is high (CT<27) in a wide range of metabolic tissues including pancreas, adrenal gland, thyroid, pituitary, adult and fetal heart, skeletal
15 muscle and adult and fetal liver. Also, moderate levels of expression is seen in in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. This expression pattern is consistent to that seen in panel 1.5. See panel 1.5 for further discussion of this gene.

Panel 1.3D Summary: Ag760 Expression of this gene is highest in small intestine
20 (CT = 26). The expression pattern is similar to that observed in Panel 1.5 and 1.2. See panel 1.5 for and panel 1.2 for further discussion of this gene.

Panel 2D Summary: Ag1537 Expression of this gene is highest in a kidney cancer (OD04340) sample (CT=25). Overall, this gene is widely expressed across this panel with high to moderate expression in both normal and adjacent cancer tissue. However, this gene
25 is more highly expressed in kidney cancer tissue than in adjacent normal tissue, consistent with expression pattern seen in panel 2.4. Therefore, this gene could be used to distinguish kidney cancers from normal kidney tissue. In addition, therapeutic modulation of this gene, through the use of small molecule drugs or antibodies, might be of benefit in the treatment of kidney cancer.

Panel 4.1D Summary: Ag4932 Highest expression of this gene is detected in lung
30 (CT=28.5). In addition, moderate levels of expression of this gene is also seen in endothelial

cells, basophils and normal tissues represented by colon, thymus and kidney. This gene codes for a variant of PV-1, a component of the endothelial fenestral and stomatal diaphragms. Expression of this gene is consistent with the pattern already reported for PV-1 (Stan *et al.*, 1999, Proc. Natl. Acad. Sci. USA 96:13203-13207, PMID: 10557298; Stan *et al.*, 2001, Genomics 72(3):304-13, PMID: 11401446). Antibodies raised against the PV-1 encoded by this gene could prevent transendothelial trafficking of inflammatory cells to different tissues sites and therefore have a potential use for treatment of inflammatory diseases including delayed type hypersensitivity, asthma, emphysema, rheumatoid arthritis and inflammatory bowel disease.

Moderate levels of expression of this gene is also seen in liver cirrhosis samples. Therefore, antibodies or small molecule therapeutics could reduce or inhibit fibrosis that occurs in liver cirrhosis.

Panel 4D Summary: Ag760 Expression of this gene is highest in lung and thymus (CTs=26.3). High expression of this gene is also seen in normal kidney and colon with more moderate expression in endothelial cells and basophils. Expression of this gene is consistent with the pattern seen in panel 4.1D and also, with the published report (Stan *et al.*, 1999, Proc. Natl. Acad. Sci. USA 96:13203-13207, PMID: 10557298; Stan *et al.*, 2001, Genomics 72(3):304-13, PMID: 11401446). See panel 4.1D for further discussion of this gene.

general oncology screening panel_v_2.4 Summary: Ag1537/Ag760 Two experiments with different probe and primer sets are in excellent agreement. Highest expression of this gene is seen in a kidney cancer sample (CTs=22.6-25). Significant expression of this gene is seen in melanoma, colon, lung, prostate, bladder and kidney cancer as well as normal tissue samples. Expression of this gene is higher in kidney cancer as compared to corresponding normal control samples. Therefore, expression of this gene may be used to distinguish kidney cancer from normal tissue and also as a marker to detect kidney cancer. Furthermore, therapeutic modulation of this gene or its protein product through the use of antibodies or small molecule drug may be useful in the treatment of melanoma, kidney, colon, lung and prostate cancers.

AF. CG52919-01: SEZ-6-like protein(7520500).

Expression of gene CG52919-01 was assessed using the primer-probe set Ag2806, described in Table AFA. Results of the RTQ-PCR runs are shown in Tables AFB, AFC, AFD and AFE.

Table AFA. Probe Name Ag2806

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' -gatgatgaggagaccaccacta-3'	22	835	384
Probe	TET-5' - atcatcaccaccaccatcaccacagt -3' -TAMRA	26	865	385
Reverse	5' -caggtagctgacctggtgtct-3'	21	893	386

5 **Table AFB.** CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag2806, Run 206976054	Tissue Name	Rel. Exp.(%) Ag2806, Run 206976054
AD 1 Hippo	10.4	Control (Path) 3 Temporal Ctx	4.7
AD 2 Hippo	15.1	Control (Path) 4 Temporal Ctx	32.8
AD 3 Hippo	4.1	AD 1 Occipital Ctx	11.5
AD 4 Hippo	4.6	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	0.0	AD 3 Occipital Ctx	5.5
AD 6 Hippo	19.6	AD 4 Occipital Ctx	18.9
Control 2 Hippo	25.5	AD 5 Occipital Ctx	12.1
Control 4 Hippo	12.0	AD 6 Occipital Ctx	31.6
Control (Path) 3 Hippo	0.7	Control 1 Occipital Ctx	2.9
AD 1 Temporal Ctx	7.7	Control 2 Occipital Ctx	57.8
AD 2 Temporal Ctx	12.0	Control 3 Occipital Ctx	13.5
AD 3 Temporal Ctx	11.1	Control 4 Occipital Ctx	4.0
AD 4 Temporal Ctx	19.5	Control (Path) 1 Occipital Ctx	100.0
AD 5 Inf Temporal Ctx	87.7	Control (Path) 2 Occipital Ctx	13.8

AD 5 SupTemporal Ctx	52.9	Control (Path) 3 Occipital Ctx	0.9
AD 6 Inf Temporal Ctx	16.4	Control (Path) 4 Occipital Ctx	14.8
AD 6 Sup Temporal Ctx	31.0	Control 1 Parietal Ctx	13.1
Control 1 Temporal Ctx	13.5	Control 2 Parietal Ctx	45.4
Control 2 Temporal Ctx	16.5	Control 3 Parietal Ctx	9.6
Control 3 Temporal Ctx	12.5	Control (Path) 1 Parietal Ctx	53.2
Control 4 Temporal Ctx	19.5	Control (Path) 2 Parietal Ctx	22.7
Control (Path) 1 Temporal Ctx	49.7	Control (Path) 3 Parietal Ctx	0.6
Control (Path) 2 Temporal Ctx	36.3	Control (Path) 4 Parietal Ctx	31.4

Table AFC. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag2806, Run 165519991	Tissue Name	Rel. Exp.(%) Ag2806, Run 165519991
Liver adenocarcinoma	2.5	Kidney (fetal)	4.8
Pancreas	0.0	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	0.0
Adrenal gland	0.0	Renal ca. RXF 393	0.0
Thyroid	0.0	Renal ca. ACHN	0.0
Salivary gland	1.5	Renal ca. UO-31	0.0
Pituitary gland	6.0	Renal ca. TK-10	0.0
Brain (fetal)	47.0	Liver	0.0
Brain (whole)	17.7	Liver (fetal)	0.0
Brain (amygdala)	22.4	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	100.0	Lung	0.0
Brain (hippocampus)	47.3	Lung (fetal)	2.3
Brain (substantia nigra)	6.5	Lung ca. (small cell) LX-1	0.0

Brain (thalamus)	39.5	Lung ca. (small cell) NCI-H69	11.8
Cerebral Cortex	25.0	Lung ca. (s.cell var.) SHP-77	19.9
Spinal cord	5.9	Lung ca. (large cell) NCI-H460	0.0
glio/astro U87-MG	3.5	Lung ca. (non-sm. cell) A549	0.0
glio/astro U-118-MG	6.3	Lung ca. (non-s.cell) NCI-H23	5.0
astrocytoma SW1783	0.0	Lung ca. (non-s.cell) HOP-62	0.0
neuro*; met SK-N-AS	2.3	Lung ca. (non-s.cl) NCI-H522	3.4
astrocytoma SF-539	0.0	Lung ca. (squam.) SW 900	0.7
astrocytoma SNB-75	4.1	Lung ca. (squam.) NCI-H596	84.7
glioma SNB-19	1.1	Mammary gland	0.0
glioma U251	8.0	Breast ca.* (pl.ef) MCF-7	1.5
glioma SF-295	2.0	Breast ca.* (pl.ef) MDA-MB-231	1.0
Heart (fetal)	0.0	Breast ca.* (pl.ef) T47D	0.0
Heart	0.0	Breast ca. BT-549	0.0
Skeletal muscle (fetal)	3.8	Breast ca. MDA-N	0.0
Skeletal muscle	0.0	Ovary	0.0
Bone marrow	6.7	Ovarian ca. OVCAR-3	3.8
Thymus	3.7	Ovarian ca. OVCAR-4	2.4
Spleen	5.5	Ovarian ca. OVCAR-5	0.0
Lymph node	11.4	Ovarian ca. OVCAR-8	2.0
Colorectal	2.3	Ovarian ca. IGROV-1	0.0
Stomach	2.3	Ovarian ca.* (ascites) SK-OV-3	0.0
Small intestine	8.7	Uterus	2.9
Colon ca. SW480	0.0	Placenta	3.6

Colon ca.* SW620(SW480 met)	0.0	Prostate	3.5
Colon ca. HT29	1.8	Prostate ca.* (bone met)PC-3	0.0
Colon ca. HCT-116	0.0	Testis	3.8
Colon ca. CaCo-2	1.8	Melanoma Hs688(A).T	0.0
Colon ca. tissue(ODO3866)	2.2	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	5.1	Melanoma UACC-62	0.0
Gastric ca.* (liver met) NCI-N87	0.3	Melanoma M14	3.8
Bladder	5.0	Melanoma LOX IMVI	0.0
Trachea	3.4	Melanoma* (met) SK- MEL-5	0.0
Kidney	0.0	Adipose	3.2

Table AFD. Panel 2D

Tissue Name	Rel. Exp.(%) Ag2806, Run 163577806	Tissue Name	Rel. Exp.(%) Ag2806, Run 163577806
Normal Colon	1.2	Kidney Margin 8120608	0.2
CC Well to Mod Diff (ODO3866)	0.0	Kidney Cancer 8120613	0.0
CC Margin (ODO3866)	0.0	Kidney Margin 8120614	0.3
CC Gr.2 rectosigmoid (ODO3868)	0.1	Kidney Cancer 9010320	0.4
CC Margin (ODO3868)	0.1	Kidney Margin 9010321	0.2
CC Mod Diff (ODO3920)	0.3	Normal Uterus	0.4
CC Margin (ODO3920)	0.4	Uterus Cancer 064011	0.7
CC Gr.2 ascend colon (ODO3921)	0.5	Normal Thyroid	0.1
CC Margin (ODO3921)	0.1	Thyroid Cancer 064010	0.0
CC from Partial Hepatectomy (ODO4309) Mets	0.3	Thyroid Cancer A302152	0.2
Liver Margin (ODO4309)	0.0	Thyroid Margin A302153	0.1

Colon mets to lung (OD04451-01)	0.2	Normal Breast	0.4
Lung Margin (OD04451-02)	0.2	Breast Cancer (OD04566)	100.0
Normal Prostate 6546-1	1.4	Breast Cancer (OD04590-01)	0.2
Prostate Cancer (OD04410)	0.5	Breast Cancer Mets (OD04590-03)	0.3
Prostate Margin (OD04410)	0.9	Breast Cancer Metastasis (OD04655-05)	0.1
Prostate Cancer (OD04720-01)	1.0	Breast Cancer 064006	0.1
Prostate Margin (OD04720-02)	0.6	Breast Cancer 1024	0.7
Normal Lung 061010	0.9	Breast Cancer 9100266	0.1
Lung Met to Muscle (ODO4286)	0.0	Breast Margin 9100265	0.1
Muscle Margin (ODO4286)	0.2	Breast Cancer A209073	0.3
Lung Malignant Cancer (OD03126)	0.1	Breast Margin A209073	0.3
Lung Margin (OD03126)	0.3	Normal Liver	0.1
Lung Cancer (OD04404)	0.1	Liver Cancer 064003	0.0
Lung Margin (OD04404)	0.3	Liver Cancer 1025	0.1
Lung Cancer (OD04565)	0.1	Liver Cancer 1026	0.3
Lung Margin (OD04565)	0.3	Liver Cancer 6004-T	0.1
Lung Cancer (OD04237-01)	0.2	Liver Tissue 6004-N	0.2
Lung Margin (OD04237-02)	0.2	Liver Cancer 6005-T	0.2
Ocular Mel Met to Liver (ODO4310)	0.1	Liver Tissue 6005-N	0.0
Liver Margin (ODO4310)	0.0	Normal Bladder	0.2
Melanoma Mets to Lung (OD04321)	0.0	Bladder Cancer 1023	0.2
Lung Margin (OD04321)	0.4	Bladder Cancer A302173	0.2
Normal Kidney	0.3	Bladder Cancer (OD04718-01)	0.1

Kidney Ca, Nuclear grade 2 (OD04338)	0.4	Bladder Normal Adjacent (OD04718-03)	0.5
Kidney Margin (OD04338)	85.3	Normal Ovary	0.1
Kidney Ca Nuclear grade 1/2 (OD04339)	0.2	Ovarian Cancer 064008	0.2
Kidney Margin (OD04339)	0.2	Ovarian Cancer (OD04768-07)	0.1
Kidney Ca, Clear cell type (OD04340)	0.1	Ovary Margin (OD04768-08)	0.1
Kidney Margin (OD04340)	0.2	Normal Stomach	0.8
Kidney Ca, Nuclear grade 3 (OD04348)	0.0	Gastric Cancer 9060358	0.1
Kidney Margin (OD04348)	0.3	Stomach Margin 9060359	0.1
Kidney Cancer (OD04622-01)	0.1	Gastric Cancer 9060395	0.3
Kidney Margin (OD04622-03)	0.0	Stomach Margin 9060394	0.5
Kidney Cancer (OD04450-01)	0.2	Gastric Cancer 9060397	0.1
Kidney Margin (OD04450-03)	0.2	Stomach Margin 9060396	0.1
Kidney Cancer 8120607	0.0	Gastric Cancer 064005	0.2

Table AFE. Panel 4D

Tissue Name	Rel. Exp.(%) Ag2806, Run 162330998	Tissue Name	Rel. Exp.(%) Ag2806, Run 162330998
Secondary Th1 act	20.3	HUVEC IL-1beta	0.0
Secondary Th2 act	5.1	HUVEC IFN gamma	21.9
Secondary Tr1 act	9.7	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	21.9	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	34.6	HUVEC IL-11	5.6
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	11.2	Lung Microvascular EC TNFalpha + IL-1beta	0.0

Primary Th2 act	13.2	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	12.8
Primary Th1 rest	37.6	Bronchial epithelium TNFalpha + IL1beta	12.1
Primary Th2 rest	10.0	Small airway epithelium none	0.0
Primary Tr1 rest	2.3	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	11.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	11.8	Coronary artery SMC TNFalpha + IL-1beta	4.4
CD8 lymphocyte act	46.7	Astrocytes rest	12.8
Secondary CD8 lymphocyte rest	44.4	Astrocytes TNFalpha + IL-1beta	11.3
Secondary CD8 lymphocyte act	10.2	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	43.2	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	54.7	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	13.2	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	17.1	Liver cirrhosis	96.6
LAK cells IL-2+IL-12	17.4	Lupus kidney	0.0
LAK cells IL-2+IFN gamma	70.2	NCI-H292 none	22.8
LAK cells IL-2+ IL-18	2.5	NCI-H292 IL-4	0.0
LAK cells PMA/ionomycin	5.0	NCI-H292 IL-9	23.2
NK Cells IL-2 rest	21.0	NCI-H292 IL-13	32.1
Two Way MLR 3 day	43.5	NCI-H292 IFN gamma	11.4
Two Way MLR 5 day	18.0	HPAEC none	0.0
Two Way MLR 7 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
PBMC rest	7.2	Lung fibroblast none	0.0

PBMC PWM	48.3	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PHA-L	76.8	Lung fibroblast IL-4	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-9	12.2
Ramos (B cell) ionomycin	0.0	Lung fibroblast IL-13	0.0
B lymphocytes PWM	28.7	Lung fibroblast IFN gamma	0.0
B lymphocytes CD40L and IL-4	100.0	Dermal fibroblast CCD1070 rest	0.0
EOL-1 dbcAMP	21.2	Dermal fibroblast CCD1070 TNF alpha	85.3
EOL-1 dbcAMP PMA/ionomycin	50.7	Dermal fibroblast CCD1070 IL-1 beta	0.0
Dendritic cells none	15.3	Dermal fibroblast IFN gamma	0.0
Dendritic cells LPS	32.3	Dermal fibroblast IL-4	0.0
Dendritic cells anti-CD40	31.4	IBD Colitis 2	12.1
Monocytes rest	52.5	IBD Crohn's	5.6
Monocytes LPS	42.3	Colon	94.0
Macrophages rest	11.0	Lung	27.7
Macrophages LPS	0.0	Thymus	34.6
HUVEC none	0.0	Kidney	78.5
HUVEC starved	11.7		

CNS_neurodegeneration_v1.0 Summary: Ag2806 This panel confirms the expression of this gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. See Panel 1.3D for a discussion of this gene in treatment of central nervous system disorders.

Panel 1.3D Summary: Ag2806 Highest expression of this gene is detected in brain cerebellum (CT=31.2). Moderate levels of expression of this gene is mainly seen in all the regions of brain including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, therapeutic modulation of this gene product may be useful in the treatment of central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

This gene codes for a homolog of mouse seizure related protein, SEZ-6. Mouse SEZ-6 was first isolated from cerebrum cortex-derived cells treated with pentylentetrazole (PTZ), one of the convulsant drugs (Shimizu-Nishikawa *et al.*, 1995, Brain Res Mol Brain Res 28(2):201-10, PMID: 7723619). Thus, SEZ-6 protein encoded by this gene may also
 5 play a role in brain seizure.

In addition, moderate to low levels of expression of this gene is also seen in three lung cancer cell lines. Therefore, expression of this gene may be used as diagnostic marker to detect lung cancer and also, modulation of this gene or its protein product through the use of antibody or protein therapeutics, may be useful in the treatment of lung cancer.

10 **Panel 2D Summary:** Ag2806 Highest expression of this gene is detected in breast cancer and normal kidney (CTs=26). Low levels of expression of this gene is also seen in breast, prostate, colon, uterine and kidney cancer. Therefore, therapeutic modulation of this gene product through the use of antibodies may be useful in the treatment of these cancers.

Panel 4D Summary: Ag2806 Highest expression of this gene is detected in CD40L
 15 and IL-4 treated B lymphocytes (CT=34). Low but significant levels of expression of this gene is also seen in TNF alpha treated dermal fibroblasts, IL-2+IFN gamma treated LAK cells, PHA-L treated PBMC cells, liver cirrhosis and normal tissue represented by colon and kidney. Therefore, therapeutic modulation of this gene may be useful in the treatment of autoimmune and inflammatory diseases such as lupus erythematosus, Crohn's disease,
 20 ulcerative colitis, multiple sclerosis, chronic obstructive pulmonary disease, asthma, emphysema, rheumatoid arthritis, or psoriasis and liver cirrhosis.

AG. , CG52919-02, CG52919-03 and CG52919-04: SEZ-6-like protein (7520500-54-1).

Expression of gene CG52919-02, CG52919-03 and CG52919-04 was assessed using
 25 the primer-probe sets Ag2795, Ag2807, Ag90 and Ag7017, described in Tables AGA, AGB, AGC and AGD. Results of the RTQ-PCR runs are shown in Tables AGE, AGF, AGG, AGH, AGI, AGJ, AGK, AGL, AGM and AGN.

Table AGA. Probe Name Ag2795

Primers	Sequences	Length	Start Position	SEQ ID No
---------	-----------	--------	----------------	-----------

Forward	5'-cctacaaccgcattaccataga-3'	22	1670	387
Probe	TET-5'-tcagcggttgacaatccaacttacga-3'-TAMRA	26	1693	388
Reverse	5'-cccacctagatggagacttcac-3'	22	1739	389

Table AGB. Probe Name Ag2807

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-cctacaaccgcattaccataga-3'	22	1670	390
Probe	TET-5'-tcagcggttgacaatccaacttacga-3'-TAMRA	26	1693	391
Reverse	5'-cccacctagatggagacttcac-3'	22	1739	392

Table AGC. Probe Name Ag90

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ttggcctggactgcttcttc-3'	20	977	393
Probe	TET-5'-catctctgtctaccctggctatggcgtg-3'-TAMRA	28	999	394
Reverse	5'-aggetgatattctggaccttgatt-3'	24	1029	395

Table AGD. Probe Name Ag7017

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gtttgacaatccaacttacgagac-3'	24	1698	396
Probe	TET-5'-cctagatggagacttcacattctctcgtc-3'-TAMRA	30	1727	397
Reverse	5'-caagtctgagttgacttccttagac-3'	25	1765	398

5

Table AGE. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag2795, Run 255324382	Tissue Name	Rel. Exp.(%) Ag2795, Run 255324382
110967 COPD-F	0.0	112427 Match Control Psoriasis-F	0.1
110980 COPD-F	0.0	112418 Psoriasis-M	0.0
110968 COPD-M	0.0	112723 Match Control Psoriasis-M	0.0
110977 COPD-M	0.1	112419 Psoriasis-M	0.0
110989 Emphysema-F	0.0	112424 Match Control Psoriasis-M	0.1
110992 Emphysema-F	0.0	112420 Psoriasis-M	0.1
110993 Emphysema-F	0.0	112425 Match Control Psoriasis-M	0.1
110994 Emphysema-F	0.0	104689 (MF) OA Bone- Backus	0.1
110995 Emphysema-F	0.2	104690 (MF) Adj "Normal" Bone-Backus	0.3
110996 Emphysema-F	0.0	104691 (MF) OA Synovium-Backus	0.0
110997 Asthma-M	0.0	104692 (BA) OA Cartilage-Backus	0.0
111001 Asthma-F	0.1	104694 (BA) OA Bone- Backus	0.1
111002 Asthma-F	0.0	104695 (BA) Adj "Normal" Bone-Backus	0.0
111003 Atopic Asthma- F	0.0	104696 (BA) OA Synovium-Backus	0.0
111004 Atopic Asthma- F	0.0	104700 (SS) OA Bone- Backus	0.0
111005 Atopic Asthma- F	0.0	104701 (SS) Adj "Normal" Bone-Backus	0.0
111006 Atopic Asthma- F	0.0	104702 (SS) OA Synovium-Backus	0.0
111417 Allergy-M	0.0	117093 OA Cartilage Rep7	0.0
112347 Allergy-M	0.0	112672 OA Bone5	0.0
112349 Normal Lung-F	0.0	112673 OA Synovium5	0.0

I12357 Normal Lung-F	0.1	I12674 OA Synovial Fluid cells5	0.0
I12354 Normal Lung-M	0.0	I17100 OA Cartilage Rep14	0.0
I12374 Crohns-F	0.0	I12756 OA Bone9	100.0
I12389 Match Control Crohns-F	0.0	I12757 OA Synovium9	0.0
I12375 Crohns-F	0.2	I12758 OA Synovial Fluid Cells9	0.0
I12732 Match Control Crohns-F	0.0	I17125 RA Cartilage Rep2	0.1
I12725 Crohns-M	0.0	I13492 Bone2 RA	0.0
I12387 Match Control Crohns-M	0.0	I13493 Synovium2 RA	0.0
I12378 Crohns-M	0.0	I13494 Syn Fluid Cells RA	0.0
I12390 Match Control Crohns-M	0.2	I13499 Cartilage4 RA	0.0
I12726 Crohns-M	0.1	I13500 Bone4 RA	0.2
I12731 Match Control Crohns-M	0.1	I13501 Synovium4 RA	0.0
I12380 Ulcer Col-F	0.1	I13502 Syn Fluid Cells4 RA	0.0
I12734 Match Control Ulcer Col-F	0.1	I13495 Cartilage3 RA	0.0
I12384 Ulcer Col-F	0.4	I13496 Bone3 RA	0.3
I12737 Match Control Ulcer Col-F	0.0	I13497 Synovium3 RA	0.2
I12386 Ulcer Col-F	0.0	I13498 Syn Fluid Cells3 RA	0.0
I12738 Match Control Ulcer Col-F	0.0	I17106 Normal Cartilage Rep20	0.0
I12381 Ulcer Col-M	0.0	I13663 Bone3 Normal	0.0
I12735 Match Control Ulcer Col-M	0.7	I13664 Synovium3 Normal	0.0
I12382 Ulcer Col-M	0.0	I13665 Syn Fluid Cells3 Normal	0.0

112394 Match Control Ulcer Col-M	0.0	117107 Normal Cartilage Rep22	0.0
112383 Ulcer Col-M	0.1	113667 Bone4 Normal	0.0
112736 Match Control Ulcer Col-M	0.0	113668 Synovium4 Normal	0.0
112423 Psoriasis-F	0.1	113669 Syn Fluid Cells4 Normal	0.0

Table AGE. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag2795, Run 20697605 2	Rel. Exp.(%) Ag2807, Run 20648228 2	Rel. Exp.(%) Ag7017, Run 279032451	Tissue Name	Rel. Exp.(%) Ag2795, Run 206976052	Rel. Exp.(%) Ag2807, Run 206482282	Rel. Exp.(%) Ag7017, Run 27903245 1
AD 1 Hippo	13.4	14.6	10.6	Control (Path) 3 Temporal Ctx	3.5	2.4	3.9
AD 2 Hippo	66.4	80.1	62.4	Control (Path) 4 Temporal Ctx	25.2	28.1	23.3
AD 3 Hippo	7.5	6.0	7.3	AD 1 Occipital Ctx	6.7	7.9	9.3
AD 4 Hippo	7.4	11.1	9.5	AD 2 Occipital Ctx (Missing)	0.0	0.0	0.0
AD 5 hippo	0.0	77.4	61.6	AD 3 Occipital Ctx	1.8	1.7	2.6
AD 6 Hippo	53.6	49.7	46.0	AD 4 Occipital Ctx	19.9	20.9	21.0
Control 2 Hippo	49.7	48.3	51.8	AD 5 Occipital Ctx	7.9	9.8	5.6
Control 4 Hippo	6.6	8.1	7.3	AD 6 Occipital Ctx	54.3	66.4	50.3
Control (Path) 3 Hippo	2.0	2.4	0.0	Control 1 Occipital Ctx	1.3	2.2	1.8

AD 1 Temporal Ctx	6.8	6.9	9.3	Control 2 Occipital Ctx	60.7	69.3	97.3
AD 2 Temporal Ctx	27.9	34.4	25.7	Control 3 Occipital Ctx	14.2	16.3	13.4
AD 3 Temporal Ctx	4.1	2.3	6.5	Control 4 Occipital Ctx	3.4	2.9	3.4
AD 4 Temporal Ctx	19.1	20.7	21.2	Control (Path) 1 Occipital Ctx	100.0	100.0	79.6
AD 5 Inf Temporal Ctx	59.5	69.3	100.0	Control (Path) 2 Occipital Ctx	8.5	7.6	6.6
AD 5 SupTemporal Ctx	46.0	45.7	39.5	Control (Path) 3 Occipital Ctx	0.7	1.2	1.7
AD 6 Inf Temporal Ctx	18.9	26.2	21.5	Control (Path) 4 Occipital Ctx	10.0	11.0	10.6
AD 6 Sup Temporal Ctx	20.9	21.2	22.2	Control 1 Parietal Ctx	6.4	5.0	3.5
Control 1 Temporal Ctx	3.1	2.5	3.9	Control 2 Parietal Ctx	22.8	26.2	24.7
Control 2 Temporal Ctx	48.3	56.6	52.5	Control 3 Parietal Ctx	14.9	20.0	20.4
Control 3 Temporal Ctx	12.6	15.4	14.7	Control (Path) 1 Parietal Ctx	74.2	92.7	84.7
Control 4 Temporal Ctx	5.9	8.5	6.7	Control (Path) 2 Parietal Ctx	20.3	20.4	20.0
Control (Path) 1 Temporal Ctx	74.7	81.2	71.2	Control (Path) 3 Parietal Ctx	1.4	1.3	2.9
Control (Path) 2 Temporal Ctx	29.1	40.1	30.4	Control (Path) 4 Parietal Ctx	30.8	40.1	38.2

Table AGG. General_screening_panel_v1.6

Tissue Name	Rel. Exp.(%) Ag7017, Run 279032445	Tissue Name	Rel. Exp.(%) Ag7017, Run 279032445
Adipose	0.1	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	0.0
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.1
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.1
Testis Pool	0.1	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.0
Prostate Pool	0.0	Colon ca. CaCo-2	0.0
Placenta	0.2	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.1	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.0
Ovarian ca. OVCAR-5	0.1	Small Intestine Pool	0.1
Ovarian ca. IGROV-1	0.8	Stomach Pool	0.0
Ovarian ca. OVCAR-8	0.3	Bone Marrow Pool	0.0
Ovary	0.0	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.0
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.0
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	0.0
Breast Pool	0.0	Thymus Pool	0.0

Trachea	0.0	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	0.1	CNS cancer (neuro;met) SK-N-AS	0.2
Lung ca. NCI-N417	4.3	CNS cancer (astro) SF- 539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB- 75	0.1
Lung ca. NCI-H146	36.9	CNS cancer (glio) SNB- 19	0.6
Lung ca. SHP-77	28.5	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	9.7
Lung ca. NCI-H526	27.9	Brain (cerebellum)	84.7
Lung ca. NCI-H23	0.1	Brain (fetal)	100.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	12.8
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	11.1
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	6.4
Liver	0.0	Brain (Thalamus) Pool	19.5
Fetal Liver	0.0	Brain (whole)	22.7
Liver ca. HepG2	0.0	Spinal Cord Pool	2.1
Kidney Pool	0.0	Adrenal Gland	0.1
Fetal Kidney	0.0	Pituitary gland Pool	1.8
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.1	Pancreas Pool	0.0

Table AGH. HASS Panel v1.0

Tissue Name	Rel. Exp.(%) Ag2795, Run 268787250	Tissue Name	Rel. Exp.(%) Ag2795, Run 268787250
MCF-7 CI	0.2	U87-MG FI (B)	0.0

MCF-7 C2	0.0	U87-MG F2	0.0
MCF-7 C3	0.4	U87-MG F3	0.0
MCF-7 C4	0.0	U87-MG F4	0.2
MCF-7 C5	0.2	U87-MG F5	0.0
MCF-7 C6	0.5	U87-MG F6	0.0
MCF-7 C7	0.4	U87-MG F7	0.1
MCF-7 C9	0.2	U87-MG F8	0.1
MCF-7 C10	0.1	U87-MG F9	0.0
MCF-7 C11	0.0	U87-MG F10	0.0
MCF-7 C12	0.2	U87-MG F11	0.0
MCF-7 C13	0.3	U87-MG F12	0.0
MCF-7 C15	0.2	U87-MG F13	0.0
MCF-7 C16	0.3	U87-MG F14	0.0
MCF-7 C17	0.4	U87-MG F15	0.1
T24 D1	0.0	U87-MG F16	0.0
T24 D2	0.1	U87-MG F17	0.0
T24 D3	0.0	LnCAP A1	0.2
T24 D4	0.0	LnCAP A2	0.5
T24 D5	0.1	LnCAP A3	0.2
T24 D6	0.0	LnCAP A4	0.9
T24 D7	0.0	LnCAP A5	0.1
T24 D9	0.0	LnCAP A6	0.3
T24 D10	0.0	LnCAP A7	0.4
T24 D11	0.0	LnCAP A8	0.2
T24 D12	0.0	LnCAP A9	0.0
T24 D13	0.0	LnCAP A10	0.3
T24 D15	0.0	LnCAP A11	0.9
T24 D16	0.0	LnCAP A12	0.0
T24 D17	0.0	LnCAP A13	0.0
CAPaN B1	0.0	LnCAP A14	0.0
CAPaN B2	0.0	LnCAP A15	0.1

CAPaN B3	0.0	LnCAP A16	0.9
CAPaN B4	0.0	LnCAP A17	0.2
CAPaN B5	0.0	Primary Astrocytes	1.0
CAPaN B6	0.0	Primary Renal Proximal Tubule Epithelial cell A2	0.0
CAPaN B7	0.0	Primary melanocytes A5	0.0
CAPaN B8	0.0	126443 - 341 medullo	0.0
CAPaN B9	0.1	126444 - 487 medullo	0.1
CAPaN B10	0.0	126445 - 425 medullo	0.5
CAPaN B11	0.1	126446 - 690 medullo	100.0
CAPaN B12	0.0	126447 - 54 adult glioma	0.3
CAPaN B13	0.0	126448 - 245 adult glioma	0.1
CAPaN B14	0.0	126449 - 317 adult glioma	3.7
CAPaN B15	0.0	126450 - 212 glioma	30.8
CAPaN B16	0.0	126451 - 456 glioma	50.0
CAPaN B17	0.0		

Table AGI. Oncology_cell_line_screening_panel_v3.2

Tissue Name	Rel. Exp.(%) Ag2795, Run 2714006 11	Tissue Name	Rel. Exp.(%) Ag2795 , Run 271400 611
94905_Daoy_Medulloblastoma/Cerebellum_sscDNA	0.0	94954_Ca Ski_Cervical epidermoid carcinoma (metastasis)_sscDNA	0.0
94906_TE671_Medulloblastom/Cerebellum_sscDNA	0.0	94955_ES-2_Ovarian clear cell carcinoma_sscDNA	0.0
94907_D283 Med_Medulloblastoma/Cerebellum_sscDNA	7.2	94957_Ramos/6h stim_Stimulated with PMA/ionomycin 6h_sscDNA	0.0
94908_PFSK-1_Primitive Neuroectodermal/Cerebellum_sscDNA	0.0	94958_Ramos/14h stim_Stimulated with PMA/ionomycin 14h_sscDNA	0.0
94909_XF-498_CNS_sscDNA	0.2	94962_MEG-01_Chronic myelogenous leukemia (megakaryoblast)_sscDNA	0.0

94910_SNB-78_CNS/glioma_sscDNA	0.0	94963_Raji_Burkitt's lymphoma_sscDNA	0.0
94911_SF-268_CNS/glioblastoma_sscDNA	0.0	94964_Daudi_Burkitt's lymphoma_sscDNA	0.0
94912_T98G_Glioblastoma_sscDNA	0.0	94965_U266_B-cell plasmacytoma/myeloma_sscDNA	0.0
96776_SK-N-SH_Neuroblastoma (metastasis)_sscDNA	0.6	94968_CA46_Burkitt's lymphoma_sscDNA	0.0
94913_SF-295_CNS/glioblastoma_sscDNA	0.0	94970_RL_non-Hodgkin's B-cell lymphoma_sscDNA	0.0
132565_NT2 pool_sscDNA	0.0	94972_JM1_pre-B-cell lymphoma/leukemia_sscDNA	0.0
94914_Cerebellum_sscDNA	14.9	94973_Jurkat_T cell leukemia_sscDNA	0.0
96777_Cerebellum_sscDNA	15.2	94974_TF-I_Erythroleukemia_sscDNA	0.0
94916_NCI-H292_Mucoepidermoid lung carcinoma_sscDNA	0.0	94975_HUT 78_T-cell lymphoma_sscDNA	0.0
94917_DMS-114_Small cell lung cancer_sscDNA	0.0	94977_U937_Histiocytic lymphoma_sscDNA	0.0
94918_DMS-79_Small cell lung cancer/neuroendocrine_sscDNA	100.0	94980_KU-812_Myelogenous leukemia_sscDNA	0.0
94919_NCI-H146_Small cell lung cancer/neuroendocrine_sscDNA	36.6	94981_769-P_Clear cell renal carcinoma_sscDNA	0.0
94920_NCI-H526_Small cell lung cancer/neuroendocrine_sscDNA	33.0	94983_Caki-2_Clear cell renal carcinoma_sscDNA	0.0
94921_NCI-N417_Small cell lung cancer/neuroendocrine_sscDNA	5.1	94984_SW 839_Clear cell renal carcinoma_sscDNA	0.0
94923_NCI-H82_Small cell lung cancer/neuroendocrine_sscDNA	6.9	94986_G401_Wilms' tumor_sscDNA	0.0
94924_NCI-H157_Squamous cell lung cancer (metastasis)_sscDNA	0.0	126768_293 cells_sscDNA	0.0
94925_NCI-H1155_Large cell lung cancer/neuroendocrine_sscDNA	7.3	94987_Hs766T_Pancreatic carcinoma (LN metastasis)_sscDNA	0.0
94926_NCI-H1299_Large cell lung cancer/neuroendocrine_sscDNA	0.0	94988_CAPAN-I_Pancreatic adenocarcinoma (liver metastasis)_sscDNA	0.0
94927_NCI-H727_Lung carcinoid_sscDNA	3.1	94989_SU86.86_Pancreatic carcinoma (liver metastasis)_sscDNA	0.0

94928_NCI-UMC-11_Lung carcinoid_sscDNA	7.5	94990_BxPC-3_Pancreatic adenocarcinoma_sscDNA	0.0
94929_LX-1_Small cell lung cancer_sscDNA	0.0	94991_HPAC_Pancreatic adenocarcinoma_sscDNA	0.0
94930_Colo-205_Colon cancer_sscDNA	0.0	94992_MIA PaCa-2_Pancreatic carcinoma_sscDNA	0.0
94931_KM12_Colon cancer_sscDNA	0.0	94993_CFPAC-1_Pancreatic ductal adenocarcinoma_sscDNA	0.0
94932_KM20L2_Colon cancer_sscDNA	0.0	94994_PANC-1_Pancreatic epithelioid ductal carcinoma_sscDNA	0.0
94933_NCI-H716_Colon cancer_sscDNA	3.6	94996_T24_Bladder carcinma (transitional cell)_sscDNA	0.0
94935_SW-48_Colon adenocarcinoma_sscDNA	0.0	94997_5637_Bladder carcinoma_sscDNA	0.0
94936_SW1116_Colon adenocarcinoma_sscDNA	0.0	94998_HT-1197_Bladder carcinoma_sscDNA	0.0
94937_LS 174T_Colon adenocarcinoma_sscDNA	0.0	94999_UM-UC-3_Bladder carcinma (transitional cell)_sscDNA	0.0
94938_SW-948_Colon adenocarcinoma_sscDNA	0.0	95000_A204_Rhabdomyosarcoma_sscDNA	0.0
94939_SW-480_Colon adenocarcinoma_sscDNA	0.0	95001_HT-1080_Fibrosarcoma_sscDNA	0.0
94940_NCI-SNU-5_Gastric carcinoma_sscDNA	0.0	95002_MG-63_Osteosarcoma (bone)_sscDNA	0.0
112197_KATO III_Stomach_sscDNA	0.0	95003_SK-LMS-1_Leiomyosarcoma (vulva)_sscDNA	0.0
94943_NCI-SNU-16_Gastric carcinoma_sscDNA	0.0	95004_SJRH30_Rhabdomyosarcoma (met to bone marrow)_sscDNA	0.8
94944_NCI-SNU-1_Gastric carcinoma_sscDNA	0.0	95005_A431_Epidermoid carcinoma_sscDNA	0.0
94946_RF-1_Gastric adenocarcinoma_sscDNA	0.0	95007_WM266-4_Melanoma_sscDNA	0.0
94947_RF-48_Gastric adenocarcinoma_sscDNA	0.0	112195_DU 145_Prostate_sscDNA	0.0
96778_MKN-45_Gastric carcinoma_sscDNA	0.0	95012_MDA-MB-468_Breast adenocarcinoma_sscDNA	0.0
94949_NCI-N87_Gastric carcinoma_sscDNA	0.0	112196_SSC-4_Tongue_sscDNA	0.0

94951_OVCAR-5_Ovarian carcinoma_sscDNA	0.0	112194_SSC-9_Tongue_sscDNA	0.0
94952_RL95-2_Uterine carcinoma_sscDNA	0.0	112191_SSC-15_Tongue_sscDNA	0.0
94953_HelaS3_Cervical adenocarcinoma_sscDNA	0.0	95017_CAL 27_Squamous cell carcinoma of tongue_sscDNA	0.0

Table AGJ. Panel I

Tissue Name	Rel. Exp.(%) Ag90, Run 87586258	Tissue Name	Rel. Exp.(%) Ag90, Run 87586258
Endothelial cells	0.0	Renal ca. 786-0	0.0
Endothelial cells (treated)	0.0	Renal ca. A498	0.0
Pancreas	0.1	Renal ca. RXF 393	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. ACHN	0.0
Adrenal gland	0.0	Renal ca. UO-31	0.0
Thyroid	0.0	Renal ca. TK-10	0.0
Salivary gland	0.0	Liver	0.0
Pituitary gland	0.0	Liver (fetal)	0.0
Brain (fetal)	37.1	Liver ca. (hepatoblast) HepG2	0.0
Brain (whole)	22.5	Lung	0.0
Brain (amygdala)	24.8	Lung (fetal)	0.0
Brain (cerebellum)	100.0	Lung ca. (small cell) LX-1	0.0
Brain (hippocampus)	29.5	Lung ca. (small cell) NCI- H69	33.7
Brain (substantia nigra)	7.6	Lung ca. (s.cell var.) SHP- 77	0.0
Brain (thalamus)	13.7	Lung ca. (large cell) NCI- H460	0.0
Brain (hypothalamus)	7.7	Lung ca. (non-sm. cell) A549	0.0
Spinal cord	1.4	Lung ca. (non-s.cell) NCI- H23	0.0

glio/astro U87-MG	0.0	Lung ca. (non-s.cell) HOP-62	0.0
glio/astro U-118-MG	0.0	Lung ca. (non-s.cl) NCI- H522	0.0
astrocytoma SW1783	0.0	Lung ca. (squam.) SW 900	0.0
neuro*; met SK-N-AS	0.4	Lung ca. (squam.) NCI- H596	20.0
astrocytoma SF-539	0.0	Mammary gland	0.1
astrocytoma SNB-75	0.0	Breast ca.* (pl.ef) MCF-7	0.0
glioma SNB-19	1.8	Breast ca.* (pl.ef) MDA- MB-231	0.0
glioma U251	0.4	Breast ca.* (pl. ef) T47D	0.0
glioma SF-295	0.0	Breast ca. BT-549	0.0
Heart	0.0	Breast ca. MDA-N	0.0
Skeletal muscle	0.0	Ovary	0.0
Bone marrow	0.0	Ovarian ca. OVCAR-3	0.0
Thymus	0.1	Ovarian ca. OVCAR-4	0.0
Spleen	0.0	Ovarian ca. OVCAR-5	0.0
Lymph node	0.0	Ovarian ca. OVCAR-8	0.0
Colon (ascending)	0.1	Ovarian ca. IGROV-1	0.0
Stomach	0.1	Ovarian ca. (ascites) SK- OV-3	0.0
Small intestine	0.3	Uterus	0.0
Colon ca. SW480	0.0	Placenta	0.0
Colon ca.* SW620 (SW480 met)	0.0	Prostate	0.0
Colon ca. HT29	0.0	Prostate ca.* (bone met) PC-3	0.0
Colon ca. HCT-116	0.0	Testis	1.3
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
Colon ca. HCT-15	0.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	0.0	Melanoma UACC-62	0.0
Gastric ca. * (liver met) NCI-N87	0.0	Melanoma M14	0.0

Bladder	0.0	Melanoma LOX IMVI	0.0
Trachea	0.0	Melanoma* (met) SK-MEL-5	0.0
Kidney	0.0	Melanoma SK-MEL-28	0.0
Kidney (fetal)	0.0		

Table AGK. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag2795, Run 165643063	Rel. Exp.(%) Ag2807, Run 165528058	Tissue Name	Rel. Exp.(%) Ag2795, Run 165643063	Rel. Exp.(%) Ag2807, Run 165528058
Liver adenocarcinoma	0.0	0.0	Kidney (fetal)	0.0	0.0
Pancreas	0.0	0.2	Renal ca. 786-0	0.0	0.0
Pancreatic ca. CAPAN 2	0.0	0.0	Renal ca. A498	0.3	0.3
Adrenal gland	0.0	0.3	Renal ca. RXF 393	0.0	0.0
Thyroid	0.0	0.0	Renal ca. ACHN	0.0	0.0
Salivary gland	0.0	0.0	Renal ca. UO-31	0.0	0.0
Pituitary gland	6.8	6.9	Renal ca. TK-10	0.0	0.0
Brain (fetal)	100.0	100.0	Liver	0.2	0.0
Brain (whole)	51.4	56.3	Liver (fetal)	0.0	0.0
Brain (amygdala)	77.4	78.5	Liver ca. (hepatoblast) HepG2	0.0	0.0
Brain (cerebellum)	58.6	79.0	Lung	0.0	0.0
Brain (hippocampus)	49.0	53.2	Lung (fetal)	0.1	0.2
Brain (substantia nigra)	9.7	13.5	Lung ca. (small cell) LX-1	0.0	0.0
Brain (thalamus)	46.7	63.7	Lung ca. (small cell) NCI-H69	41.5	92.7
Cerebral Cortex	37.9	32.5	Lung ca. (s.cell var.) SHP-77	34.4	25.2
Spinal cord	5.9	4.1	Lung ca. (large cell) NCI-H460	0.3	0.0

glio/astro U87-MG	0.0	0.0	Lung ca. (non-sm. cell) A549	0.0	0.0
glio/astro U-118-MG	0.0	0.0	Lung ca. (non-s.cell) NCI-H23	0.1	0.0
astrocytoma SW1783	0.0	0.1	Lung ca. (non-s.cell) HOP-62	0.0	0.0
neuro*; met SK-N-AS	0.4	0.0	Lung ca. (non-s.cl) NCI-H522	0.0	0.0
astrocytoma SF-539	0.0	0.0	Lung ca. (squam.) SW 900	0.0	0.0
astrocytoma SNB-75	0.0	0.0	Lung ca. (squam.) NCI-H596	70.2	58.2
glioma SNB-19	1.2	2.5	Mammary gland	0.4	1.1
glioma U251	3.5	4.2	Breast ca.* (pl.ef) MCF-7	0.0	0.0
glioma SF-295	0.0	0.0	Breast ca.* (pl.ef) MDA-MB-231	0.0	0.0
Heart (fetal)	0.0	0.0	Breast ca.* (pl.ef) T47D	0.0	0.0
Heart	0.1	0.0	Breast ca. BT-549	0.0	0.0
Skeletal muscle (fetal)	0.4	0.0	Breast ca. MDA-N	0.0	0.0
Skeletal muscle	0.0	0.0	Ovary	0.0	0.0
Bone marrow	0.0	0.0	Ovarian ca. OVCAR-3	0.0	0.2
Thymus	0.0	0.0	Ovarian ca. OVCAR-4	0.1	0.0
Spleen	0.5	0.0	Ovarian ca. OVCAR-5	0.0	0.0
Lymph node	0.1	0.1	Ovarian ca. OVCAR-8	0.0	0.0
Colorectal	0.0	0.0	Ovarian ca. IGROV-1	0.0	0.0
Stomach	0.0	0.0	Ovarian ca.* (ascites) SK-OV-3	0.2	0.0
Small intestine	0.6	1.0	Uterus	0.0	0.1
Colon ca. SW480	0.0	0.0	Placenta	0.1	0.2

Colon ca.* SW620(SW480 met)	0.0	0.0	Prostate	0.0	0.0
Colon ca. HT29	0.0	0.0	Prostate ca.* (bone met)PC-3	0.0	0.0
Colon ca. HCT-116	0.3	0.0	Testis	0.4	0.3
Colon ca. CaCo-2	0.0	0.0	Melanoma Hs688(A).T	0.0	0.0
Colon ca. tissue(ODO3866)	0.0	0.0	Melanoma* (met) Hs688(B).T	0.1	0.0
Colon ca. HCC-2998	0.0	0.0	Melanoma UACC- 62	0.0	0.0
Gastric ca.* (liver met) NCI-N87	0.4	0.1	Melanoma M14	0.0	0.0
Bladder	0.4	0.1	Melanoma LOX IMVI	0.0	0.0
Trachea	0.1	0.1	Melanoma* (met) SK-MEL-5	0.0	0.1
Kidney	0.0	0.0	Adipose	0.1	0.0

Table AGL. Panel 2D

Tissue Name	Rel. Exp.(%) Ag2795, Run 163577802	Rel. Exp.(%) Ag2807, Run 162598819	Tissue Name	Rel. Exp.(%) Ag2795, Run 163577802	Rel. Exp.(%) Ag2807, Run 162598819
Normal Colon	6.3	13.2	Kidney Margin 8120608	0.0	3.8
CC Well to Mod Diff (ODO3866)	0.0	0.0	Kidney Cancer 8120613	0.5	0.0
CC Margin (ODO3866)	1.7	5.9	Kidney Margin 8120614	2.5	1.5
CC Gr.2 rectosigmoid (ODO3868)	0.0	0.0	Kidney Cancer 9010320	0.0	0.7
CC Margin (ODO3868)	0.0	2.9	Kidney Margin 9010321	1.7	1.7
CC Mod Diff (ODO3920)	0.0	0.7	Normal Uterus	0.0	1.0
CC Margin (ODO3920)	0.6	1.4	Uterus Cancer 064011	1.5	0.8

CC Gr.2 ascend colon (ODO3921)	0.0	3.4	Normal Thyroid	1.3	0.3
CC Margin (ODO3921)	1.0	2.3	Thyroid Cancer 064010	1.6	0.0
CC from Partial Hepatectomy (ODO4309) Mets	0.0	0.4	Thyroid Cancer A302152	1.3	0.2
Liver Margin (ODO4309)	0.0	0.4	Thyroid Margin A302153	0.0	0.0
Colon mets to lung (OD04451-01)	0.0	2.9	Normal Breast	0.7	1.7
Lung Margin (OD04451-02)	1.1	4.5	Breast Cancer (OD04566)	0.0	0.3
Normal Prostate 6546-1	6.7	9.8	Breast Cancer (OD04590-01)	0.9	1.3
Prostate Cancer (OD04410)	2.3	3.4	Breast Cancer Mets (OD04590-03)	0.6	1.8
Prostate Margin (OD04410)	1.6	3.3	Breast Cancer Metastasis (OD04655-05)	1.0	0.8
Prostate Cancer (OD04720-01)	1.2	2.6	Breast Cancer 064006	0.0	0.0
Prostate Margin (OD04720-02)	9.0	7.2	Breast Cancer 1024	0.0	1.5
Normal Lung 061010	2.1	2.6	Breast Cancer 9100266	0.3	1.2
Lung Met to Muscle (ODO4286)	0.0	0.0	Breast Margin 9100265	0.0	1.7
Muscle Margin (ODO4286)	0.7	0.0	Breast Cancer A209073	0.3	0.0
Lung Malignant Cancer (OD03126)	0.0	0.0	Breast Margin A209073	0.0	0.0
Lung Margin (OD03126)	0.5	0.5	Normal Liver	0.0	0.4
Lung Cancer (OD04404)	0.0	0.0	Liver Cancer 064003	0.0	0.0
Lung Margin (OD04404)	0.0	1.5	Liver Cancer 1025	0.0	0.0

Lung Cancer (OD04565)	0.0	0.3	Liver Cancer 1026	100.0	100.0
Lung Margin (OD04565)	0.0	1.7	Liver Cancer 6004-T	0.0	0.0
Lung Cancer (OD04237-01)	20.4	13.2	Liver Tissue 6004-N	0.0	0.0
Lung Margin (OD04237-02)	1.3	0.9	Liver Cancer 6005-T	80.1	64.6
Ocular Mel Met to Liver (ODO4310)	0.0	2.6	Liver Tissue 6005-N	0.0	0.0
Liver Margin (ODO4310)	0.0	0.0	Normal Bladder	7.2	7.4
Melanoma Mets to Lung (OD04321)	0.0	0.0	Bladder Cancer 1023	0.4	0.0
Lung Margin (OD04321)	1.6	2.9	Bladder Cancer A302173	1.3	2.8
Normal Kidney	1.1	1.0	Bladder Cancer (OD04718-01)	0.0	0.4
Kidney Ca, Nuclear grade 2 (OD04338)	1.2	1.2	Bladder Normal Adjacent (OD04718-03)	0.7	0.4
Kidney Margin (OD04338)	0.5	1.1	Normal Ovary	0.0	0.0
Kidney Ca Nuclear grade 1/2 (OD04339)	0.0	0.5	Ovarian Cancer 064008	0.0	0.7
Kidney Margin (OD04339)	0.5	0.9	Ovarian Cancer (OD04768-07)	0.4	1.1
Kidney Ca, Clear cell type (OD04340)	0.0	0.5	Ovary Margin (OD04768-08)	0.0	0.0
Kidney Margin (OD04340)	1.0	0.3	Normal Stomach	1.2	2.6
Kidney Ca, Nuclear grade 3 (OD04348)	0.0	0.9	Gastric Cancer 9060358	0.0	0.0
Kidney Margin (OD04348)	0.3	0.8	Stomach Margin 9060359	0.4	1.4
Kidney Cancer (OD04622-01)	0.0	0.9	Gastric Cancer 9060395	0.4	0.5
Kidney Margin (OD04622-03)	0.0	0.0	Stomach Margin 9060394	0.6	2.0

Kidney Cancer (OD04450-01)	0.0	1.2	Gastric Cancer 9060397	1.4	0.8
Kidney Margin (OD04450-03)	0.3	0.4	Stomach Margin 9060396	6.0	1.9
Kidney Cancer 8120607	0.0	0.0	Gastric Cancer 064005	3.4	1.6

Table AGM. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag7017, Run 279031713	Tissue Name	Rel. Exp.(%) Ag7017, Run 279031713
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1 beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1 beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1 beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1 beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	2.1	Coronary artery SMC TNFalpha + IL-1 beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	42.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL- 1 beta	15.5

Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	5.3
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	5.5	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	13.6	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0

Dendritic cells anti-CD40	0.0	Neutrophils TNF α +LPS	100.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	14.3	Colon	6.8
Macrophages rest	0.0	Lung	0.0
Macrophages LPS	0.0	Thymus	0.0
HUVEC none	0.0	Kidney	4.1
HUVEC starved	0.0		

Table AGN. Panel 4D

Tissue Name	Rel. Exp.(%) Ag2807, Run 165806333	Tissue Name	Rel. Exp.(%) Ag2807, Run 165806333
Secondary Th1 act	1.2	HUVEC IL-1 β	0.0
Secondary Th2 act	1.1	HUVEC IFN γ	2.1
Secondary Tr1 act	2.7	HUVEC TNF α + IFN γ	0.0
Secondary Th1 rest	1.9	HUVEC TNF α + IL4	0.0
Secondary Th2 rest	2.4	HUVEC IL-11	0.9
Secondary Tr1 rest	0.0	Lung Microvascular EC none	1.0
Primary Th1 act	0.0	Lung Microvascular EC TNF α + IL-1 β	0.0
Primary Th2 act	1.2	Microvascular Dermal EC none	1.9
Primary Tr1 act	0.0	Microvascular Dermal EC TNF α + IL-1 β	0.0
Primary Th1 rest	6.1	Bronchial epithelium TNF α + IL1 β	0.0
Primary Th2 rest	2.7	Small airway epithelium none	3.9
Primary Tr1 rest	1.1	Small airway epithelium TNF α + IL-1 β	3.6
CD45RA CD4 lymphocyte act	4.0	Coronary artery SMC rest	0.8
CD45RO CD4 lymphocyte act	4.5	Coronary artery SMC TNF α + IL-1 β	0.0
CD8 lymphocyte act	3.2	Astrocytes rest	100.0

Secondary CD8 lymphocyte rest	2.1	Astrocytes TNFalpha + IL-1 beta	84.1
Secondary CD8 lymphocyte act	1.6	KU-812 (Basophil) rest	4.6
CD4 lymphocyte none	1.2	KU-812 (Basophil) PMA/ionomycin	1.1
2ry Th1/Th2/Tr1_anti-CD95 CH11	5.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1 beta	0.0
LAK cells IL-2	4.1	Liver cirrhosis	1.1
LAK cells IL-2+IL-12	2.6	Lupus kidney	1.6
LAK cells IL-2+IFN gamma	3.8	NCI-H292 none	6.1
LAK cells IL-2+ IL-18	1.0	NCI-H292 IL-4	2.5
LAK cells PMA/ionomycin	1.2	NCI-H292 IL-9	2.1
NK Cells IL-2 rest	2.1	NCI-H292 IL-13	2.8
Two Way MLR 3 day	3.8	NCI-H292 IFN gamma	1.0
Two Way MLR 5 day	2.6	HPAEC none	0.0
Two Way MLR 7 day	3.6	HPAEC TNF alpha + IL-1 beta	0.0
PBMC rest	1.9	Lung fibroblast none	1.3
PBMC PWM	1.3	Lung fibroblast TNF alpha + IL-1 beta	1.6
PBMC PHA-L	0.0	Lung fibroblast IL-4	2.5
Ramos (B cell) none	6.0	Lung fibroblast IL-9	2.8
Ramos (B cell) ionomycin	0.0	Lung fibroblast IL-13	1.3
B lymphocytes PWM	0.9	Lung fibroblast IFN gamma	0.0
B lymphocytes CD40L and IL-4	9.0	Dermal fibroblast CCD1070 rest	2.1
EOL-1 dbcAMP	5.6	Dermal fibroblast CCD1070 TNF alpha	4.4
EOL-1 dbcAMP PMA/ionomycin	14.8	Dermal fibroblast CCD1070 IL-1 beta	0.8
Dendritic cells none	4.3	Dermal fibroblast IFN gamma	1.5

Dendritic cells LPS	5.0	Dermal fibroblast IL-4	0.0
Dendritic cells anti-CD40	5.1	IBD Colitis 2	6.3
Monocytes rest	5.4	IBD Crohn's	0.0
Monocytes LPS	4.4	Colon	58.2
Macrophages rest	1.6	Lung	3.4
Macrophages LPS	0.0	Thymus	8.8
HUVEC none	1.2	Kidney	31.6
HUVEC starved	0.9		

AI_comprehensive_panel_v1.0 Summary: Ag2795 High expression of this gene is mostly restricted to orthoarthritis (OA) bone (CT=28). Thus, expression of this gene may be used to distinguish OA bone from other samples used in this panel. In addition, therapeutic modulation of this gene product may be useful in the treatment of orthoarthritis.

- 5 **CNS_neurodegeneration_v1.0 Summary:** Ag2795/Ag2807/Ag7017 Three experiments with two different probes and primer sets are in very good agreement. This panel confirms the expression of this gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this
- 10 experiment. See Panel 1.3D for a discussion of this gene in treatment of central nervous system disorders.

- General_screening_panel_v1.6 Summary:** Ag7017 Highest expression of this gene is detected in brain cerebellum (CT=25.3). High to moderate levels of expression of this gene is mainly seen in all the regions of brain including amygdala, hippocampus,
- 15 substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, therapeutic modulation of this gene product may be useful in the treatment of central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

- This gene codes for a homolog of mouse seizure related protein, SEZ-6. Mouse
- 20 SEZ-6 was first isolated from cerebrum cortex-derived cells treated with pentylentetrazole (PTZ), one of the convulsant drugs (Shimizu-Nishikawa *et al.*, 1995, Brain Res Mol Brain Res 28(2):201-10, PMID: 7723619). Thus, SEZ-6 protein encoded by this gene may also play a role in brain seizure.

In addition, moderate to low levels of expression of this gene is also seen in four lung cancer cell lines and a ovarian cancer cell line. Therefore, expression of this gene may be used as diagnostic marker to detect lung cancer and also, modulation of this gene or its protein product through the use of antibody or protein therapeutics, may be useful in the treatment of lung and ovarian cancers.

HASS Panel v1.0 Summary: Ag7017 Highest expression of this gene is detected in a medulloblastoma (CT=28). In addition, moderate levels of expression of this gene is also seen in glioma samples. Therefore, therapeutic modulation of this gene may be useful in the treatment of brain cancer.

Oncology_cell_line_screening_panel_v3.2 Summary: Ag2795 Highest expression of this gene is detected in small lung cancer DMS-79 cell line (CT=26.5). Moderate to low levels of expression of this gene is also seen in number of cell lines derived from lung, colon, bone and brain cancers. Therefore, expression of this gene may be used as marker to detect these cancers. In addition, therapeutic modulation of this gene through the use of antibodies or small molecule drug may be useful in the treatment of lung, colon, bone and brain cancers.

Panel I Summary: Ag90 Highest expression of this gene is detected in brain cerebellum (CT=25). High levels of expression of this gene is mainly seen in all the regions of brain including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. In addition, moderate levels of expression of this gene is also seen in two lung cancer cell lines and a glioma cell line. See panel 1.3D for further discussion of this gene.

Panel 1.3D Summary: Ag2795/Ag2807 Two experiments with same probe and primer sets are in excellent agreement with highest expression of this gene detected fetal brain (CTs=27-28.5). Moderate levels of expression of this gene is mainly seen in all the regions of brain including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, therapeutic modulation of this gene product may be useful in the treatment of central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

This gene codes for a homolog of mouse seizure related protein, SEZ-6. Mouse SEZ-6 was first isolated from cerebrum cortex-derived cells treated with pentylentetrazole